

海外优秀数学类教材系列丛书



翻译版

A History of Mathematics
An Introduction
(Second Edition)

数学史通论（第2版）

□ VICTOR J. KATZ
□ 李文林　邹建成　胥鸣伟　等译
□ 胥鸣伟　李文林　校





高等教育出版社
Higher Education Press



海外优秀数学类教材系列丛书

全书共分四大部分：6世纪前的数学；中世纪的数学（500–1000）；早期近代数学（1400–1700）；近代数学（1700–2000）。本书主要特色如下：

1．灵活的编排：尽管本书主要是按年代顺序编排的，但每一时期则是围绕某一专题展开的。读者通过查阅详尽的标题，就能对该时期历史的全程进行跟踪。

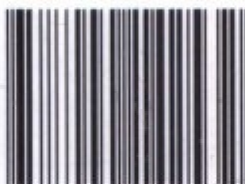
2．不同时期的重要教材：本书每一章中都会讨论一种或几种那个时期的重要教材，通过它们，不仅能学习那些伟大数学家的思想，还能看到某些论题在过去是怎样被处理的。

3．非西方数学：本书相当多的材料是关于中国、印度及伊斯兰世界的数学的；在插入章中还比较了大约在14世纪初各主要文明的数学。

4．人物传记和评注：本书配有100多张纪念历代数学家及其工作的邮票和图片，并着重用框图给出数学家的小传。

此外，本书在习题配置、专题讨论、内容的前后呼应等方面都有许多特色。本书可供综合大学、师范院校以及理工科各专业的学生作为数学史课程的教材，也可供广大数学工作者和一般科学爱好者阅读参考。相信中学师生也会从本书中获益。

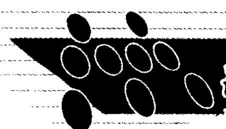
ISBN 7-04-014253-8



定价 52.00 元

PEARSON Education





海外优秀数学类教材系列丛书



翻译版

A History of Mathematics
An Introduction
(Second Edition)

数学史通论（第2版）

□ VICTOR J. KATZ

□ 李文林　邹建成　胥鸣伟
杨宝珊　刘建军　李培廉
刘向晖　吴发恩　袁　敏
王　辉　郑　权　杨浩菊　译

□ 胥鸣伟　李文林　校



高等教育出版社
Higher Education Press

图字：01－2001－2249号

VICTOR J. KATZ

A History of Mathematics: An Introduction, 2e

ISBN: 0－321－01618－1

图书在版编目（CIP）数据

数学史通论 / （美）卡茨（Katz,V.J.）著；李文林等译. —2 版. —北京：高等教育出版社，2004

书名原文：A History of Mathematics：An Introduction，2e

ISBN 7-04-014253-8

Ⅰ. 数... Ⅱ. ①卡...②李... Ⅲ. 数学史 Ⅳ. 011

中国版本图书馆 CIP 数据核字（2004）第 002142 号

责任编辑：赵天夫、徐　可　**封面设计**：王凌波　**版式设计**：杨　明　**责任印制**：陈伟光

出版发行	高等教育出版社	购书热线	010－64054588
社　址	北京市西城区德外大街4号	免费咨询	800－810－0598
邮政编码	100011	网　址	http://www.hep.edu.cn
总　机	010－82028899		http://www.hep.com.cn
经　销	新华书店北京发行所		
印　刷	涿州市星河印刷厂		
		版　次	1980年10月第1版
开　本	850×1168　1/16		2004年2月第2版
印　张	43	印　次	2004年2月第1次印刷
字　数	1140 000	定　价	52.00元

序　　言

处理方法和指导思想

美国数学协会(MAA)下属教师数学教育委员会在其《呼唤变革:关于数学教师的数学修养的建议书》中,提议所有未来中小学数学教师

> *注意培养自身对各种文化在数学思想的成长与发展过程中所作的贡献有一定的鉴赏能力;对来自各种不同文化的个人(无论男女)在古代、近代和现代数学论题的发展上所作的贡献有所研究;并对中小学数学中主要概念的历史发展有所认识.*

根据 MAA 的观点,数学史方面的知识能向学生表明,数学是一项非常重要的人类活动.数学不是一产生就像我们教科书中那样完美的形式,它常常是出于解决问题的需要,以一种直观的和实验性的形式发展出来的.数学思想的实际发展历程能有效地被用来激励和启迪今天的学生.

这本新的数学史教科书是基于这样一种认识产生的,就是,不仅未来的中小数学教师,即便是未来的大学数学教师,为了更有效地给他们的学生教好这门课,都需要对历史背景有所了解.因此这本书是为那些主修数学、今后打算在大学或高中任教的低年级或高年级的学生设计的,所以内容集中于中小学或大学本科教学计划中通常包含的那些数学课题的历史.因为一门数学课题的历史会为讲解这一课题提供非常好的思路,为了使未来的数学教师能在历史的基础上开展课堂教学,我们会对每一个新概念作充分细致的解说.实际上,许多习题就是要求读者去讲一堂课.我希望这些学生以及未来的教师能从本书获得一种关于数学的来龙去脉的知识,一种对数学中许多重要的概念有更深入的理解的知识.

本书特色

材料组织的灵活

尽管本书主要是按年代顺序划分成若干时期来进行组织的,但在每一时期内则是按专题来进行组织的.通过查阅详尽的细节标题,读者可以选择某一特定的专题,对其历史的全程进行跟踪.例如,想研究方程求解时,就可以研究古代埃及人和巴比伦人的方法,希腊人的几何解法,中国人的数值解法,伊斯兰人用圆锥截线求解三次方程的方法,意大利人所发现的求解三次方程和四次方程的一套算法,拉格朗日为解高次多项式方程而研究出来的一套判据,高斯在求解割圆方程方面所作的工作,以及伽罗瓦用置换来讨论求解方程的工作,这一工作我们今天称之为伽罗瓦理论.

关注教科书

全书对各个不同时期的一些重要教科书都给予了足够的重视.从事数学研究,发现新的定理和技巧是一回事.以一种使其他人也能掌握的方式来阐述这些定理和技巧则又是一回事.因此几乎在每一章中都会讨论一种或几种那个时代的重要的教科书.这是这样一些著作,学生们能够通过它们

来学习那些伟大的数学家们的思想.今天的学生将能够看到某些论题在过去是怎样被处理的,并能将这些处理方法与当今教科书中的方法加以比较,而且还能看到许多年前的学生想要解决的是什么样的问题.

天文和数学

有两章是完全用来讲数学方法的,也就是讲数学是怎样用于解决人类其他活动领域内的问题的.这两章,一章是关于希腊时期的,另一章则涉及文艺复兴时期,它们相当大的部分是讲述天文学的.事实上,在古代,数学家和天文学家常常是同一个人.要想了解希腊数学的主要内容,关键是要了解希腊人关于天体的模型,以及怎样借助这个模型用数学来得出预言.类似地,我们还会讨论哥白尼-开普勒的天体模型以及文艺复兴时期的数学家们是怎样用数学来研究它的.

非西方数学

我们还下了特别的功夫来讨论数学在世界上除欧洲以外一些地区的发展.于是有相当多的材料是有关中国、印度和伊斯兰世界的数学的.还安排了一个"插入章",在该章中比较了大约在14世纪之交时各主要文明的数学.接着讨论了在世界各地各种其它社会中的数学.读者会看到,有些数学概念在很多地方出现过,尽管也许并不是在我们西方称之为"数学"的背景中出现.

按专题分类的习题

每一章均含有许多习题,为了便于选取,这些习题都是按专题分类汇集的.有些习题只是简单的计算,有些则是填补正文中数学论证的空白.讨论题是一种无明确答案的漫谈式问题,其中有些可能要做些研究才能回答.很多这一类的问题要求学生动脑筋思考怎样利用在课堂上学到的历史材料.(大多数计算习题的答案可在书后的答案小节中找到.)有许多习题即使读者不打算做,他们也至少应该阅读一下这些题,以便对该章的内容有更全面的了解.

焦点论坛

小传 为了便于参阅,对许多我们谈到过他们工作的数学家,其小传被放在独立于正文的栏框中.特别是,尽管由于种种原因参与到数学研究中来的妇女为数不多,我们还是写了几位重要的女数学家的小传.妇女通常都是在克服了重重困难后才能成功地对数学事业作出贡献.

专题 还有一些特殊论题以加框文字的形式散见于全书.其中有这样一些条目,如:埃及人对希腊数学影响问题的讨论,托勒密著作中函数概念的讨论,各种连续概念的比较.还有这样一些加框读物,它们把重要的定义汇集在一起以便于查阅参考.

补充教学资料 每一章的开始有一段相关引语和对一桩重要数学"事件"的描述.每一章的末尾都有一张讲到过的数学家的简短的编年表,这将有助于学生组合他们的知识.每章还有一份附加了注释的参考文献,学生们从这些文献中可以获得更多的信息.最后,在前面有一张数学史的**大事年表**和一张标明了正文提到过的一些重要地点的位置的**地图**.

预备知识

学过一年微积分,具备了可资实用的知识,就足以理解本书的头十二章,以后的几章要求更多一些数学上的准备,但是各节的标题就清楚地表明了需要哪些数学知识.例如,要想充分理解第14章和第15章,就要求学生学过抽象代数.

本版更新之处

本书第一版获得了广泛友好的接受,这鼓励我保持它的基本体系和内容.然而,我仍力图在本书的内容及表述的清晰性两方面作出一系列的改进.改进的根据是许多使用本书第一版的人所提出的意见以及在新近文献中所刊载的有关数学史中的一些新发现.实际上每一小节都有一些小小的改动,较大的改动则有以下一些:有关伊斯兰传统组合学的新材料;牛顿对他的世界体系的推导,19 世纪和 20 世纪中的线性代数,以及 19 世纪中的统计概念.我力求改正史实上的全部错误,并杜绝新的错误,但对任何人指出本书还余留的错误,我将深表感谢.每章还增加了一些新的问题,其中有些比较简单.参考文献方面也尽可能作了更新.还增加了一些新的邮票作为插图.不过应当注意到,任何这种试图表现 16 世纪前数学家的邮票上的画像——别处的画像实际上也一样——都是想像的.至今还没有哪一张这种人物的画像是有可靠证据的.

课程内容的弹性

本书包括的材料远远超过了普通一学期的数学史的课程.实际上,它的材料适合一学年的课程.前一半内容是讲述直到 17 世纪末微积分发明为止的这一时期的.后半部分内容则是讲 18、19 和 20 世纪数学的.然而对于那些只有一个学期学时的教师来说,有几种使用本书的方式:第一种方式可以选前十二章中的绝大部分内容,然后简单地以微积分作结束;第二种方式是,可以选一到两个专题的全部历史.以下是可供选择的专题以及相应的章节:

方程求解:1.4, 1.8, 1.9, 2.4.3, 2.5, 5.2, 6.3, 6.4, 6.7, 6.8, 7.2, 8.3, 9.3, 9.4, 11.2, 14.2.4, 15.2

微积分思想:2.3.2, 2.3.3, 2.4.9, 3.2, 3.3, 7.2.4, 7.4.4, 8.4, 10.5, 12, 13, 16.1, 16.2, 16.3, 16.4

几何学的概念:1.5, 1.8, 2.1, 2.2, 2.4, 3.3, 3.4, 3.5, 4.3, 5.3, 7.4, 8.1, 10.1, 11.1, 11.5, 14.3, 17, 18.2

三角学、天文学和测量:1.6, 4.1, 4.2, 6.2, 6.6, 7.5, 8.1, 10.2, 10.3, 12.5.6, 13.1.3

组合学、概率论和统计学:6.8, 7.3, 8.2, 11.3, 14.1, 16.5

线性代数:1.4, 14.2.2, 14.2.4, 15.5, 17.4, 18.3.3, 18.4.7

数论:2.1.1, 2.4.7, 5.1, 11.4, 14.2.3, 15.1

近世代数:6.8, 7.2, 8.3, 9.1, 9.2, 14.2, 15.2, 15.3, 15.4, 18.3, 18.4.4, 18.4.6, 18.4.8

第三种方式,可以讲授前十章绝大部分的内容,然后再按某个专题,从后面的章节中选几个概念来讲.我们还可以给个别学生或学习小组指定阅读章节,并让他们做读书报告.

致　谢

和任何一本书一样,要不是有许多人的帮助,本书是不可能写成的.下面各位对本书的第一版颇多贡献,他们的投入继续对本书的改进发挥作用:Mancia Asher(Ithaca 学院),J. Lennart Berggren(Simon Fraten 大学),Robert Kreiser(A. A. U. P),Robert Rosenfeld(Nassau 社区大学)和 John Milcetich(哥伦比亚特区大学).

很多人对本书的第二版作了详尽的建议,尽管我没有全部采纳(这使我颇觉遗憾),我真诚地感谢他们为改进本书所提出的想法,这些人中有 Ivor Grattan-Guinness, Kim Plofker, Eleanor Robson,

Richard Askey, William Anglin, Claudia Zaslavsky, Rebekka Struik, William Ramaley, Joseph Albree, Calvin Jongsma, David Fowler, John Stillwell, Christian Thybo, Jim Tattersall, Judith Grabiner, Tony Gardiner, Ubi D'Ambrosio, Dirk Struik,和 David Rowe.我衷心地感谢所有这些人.

对书稿审阅的很多人也以他们细致深入的评论给了我很大的帮助,使本书增色不少,没有他们的帮助本书就不会是现在这个样子.

第一版的审稿人有:Duane Blumberg,西南路易斯安那大学;Walter Czarnec,Framingham 州立大学;Joseph Dauben,Herbert Lehman 学院 - CUNY;Harvey Davis,密执安州立大学;Joy Easton,西弗吉尼亚大学;Carl FitzGerald,加州大学圣地亚哥分校;Basil Gordon,加州大学洛杉矶分校;Mary Gray,美国大学;Branko Grunbaum,华盛顿大学;William Hintzman,圣地亚哥州立大学;Barnabas Hughes,加州州立大学 - Northridge;Israel Kleiner,York 大学;David E.Kullmam,迈阿密大学;Robert L.Hall,威斯康星大学,Milwaukee 分校;Richavd Marshall,东密执安大学;Jerold Mathews,衣阿华州立大学;Willard Parker,堪萨斯州立大学;Clinton M.Petty,Missouri-Columbia 大学,Howard Prouse,Mamkato 州立大学;Helmut Rohrl,加州大学圣地亚哥分校;David Wilson,佛罗里达大学;以及 Frederick Wright,北卡罗来纳大学 Chapel 分校.

第二版的审稿人有:Salvatore Anastasio,纽约州立大学,New Platy 分校;Bruce Crauder,Oklahoma 州立大学;Walter Czarnec,Framingham 州立大学;William England,密西西比州立大学;David Jabon,东华盛顿大学;Charles Jones,Ball 州立大学;Michael Lacey,印地安那大学;Harold Martin,北密执安大学;James Murdock,衣阿华州立大学;Ken Shaw,佛罗里达州立大学;Sverre Smalo,加州大学,Santa Barbara 分校;Domina Eberle Spencer,Connecticut 大学;Jimmy Woods,North Georgia College.

我还在各种论坛上与许多数学史家们交谈过,从中获益匪浅.特别是那些定期参加由美国国家历史博物院数学馆前馆长 Uta Merzbach 组织的数学史年会的数学家们一定会认得出有些观点是在那些年会上讨论过的.本书还从多年来与其他一些人的讨论获益,其中有 Charles Jones(Ball 州立大学),V.Frederick Rickey(Bowling Green 州立大学),Florence Fasanelli(MAA),Israel Kleiner(York 大学),Abe Shenitzer(York 大学),Ubiratan D'Ambrosio(Estadual de Campinas 大学),以及 Frank Swetz(Pennsylvania 州立大学).我在哥伦比亚特区大学数学史(及其他)班上的学生在澄清我的诸多观念上也给了我不少帮助.自然,我欢迎其他地方的学生和同事努力为进一步改进本书提出更多的意见和给我来信.

还要特别感谢哥伦比亚特区大学图书馆馆员们,特别是 Clement Goddard,他总是设法通过馆际交流为我找到任何我想要的那些不见经传的书籍.Smithsonian 协会图书馆特别收藏部的 Leslie Overstreet 在寻找图片来源时对我的帮助特别大.

感谢 Harper Collins 出版社的前编辑 Steve Quigley,Don Gecewicz 和 George Duda,他们帮助我完成了本书的第一版.

我还应感谢 Addison Wesley Longman 出版社本书的新编辑 Jennifer Albanese,感谢她在完成本书的出版中所作的建议和表现出的耐心.还有 Rebecca Malone 和 Barbara Pendergast,感谢她们在安排印装等生产方面的事务上所作的努力,还要感谢 Susan Holbert 在准备索引上所作的工作.

我的家庭在我撰写本书的多年中给了我极大的支持,我感谢我的父母对我的信任和耐心;感谢我的孩子 Sharon,Ari 和 Naomi,他们常常给我帮助,特别是允许我使用我们的电脑;最后我要感谢我的妻子菲丽丝,不论白天还是夜晚都和我进行长时间的讨论,在我需要时总是在我身边,我欠她的远远超过我能偿还的.

V.J.卡兹

目　　录

第二篇 中世纪的数学:500 - 1400

第三篇 早期近代数学:1400 - 1700

第四篇　近代数学:1700－2000

第1章 古代数学

> 精确计算,通向世间万物和一切奥秘的知识的大门.
>
> ——《兰德数学纸草书》引言[1]

美索不达米亚:大约3800年前在拉沙的一所培养书记员的学校里,有一位教师正在编撰一些数学问题给他的学生去做,以便他们练习一下刚刚讲过的关于直角三角形三边之间关系的结论.这位老师既要让这个计算有一定的难度,这样才能看出谁真正学懂了这些知识,又要让算出的结果为整数,以免学生的学习积极性受到挫伤.经过几个小时反复摆弄他所知道的不多几个能满足方程 $a^2 + b^2 = c^2$ 的三数组 (a,b,c) 之后,他有了一个新的想法.他用笔尖在一块潮湿的泥板上划了几下,很快做了些计算,这使他相信自己已经发现了怎样来算出这种三数组,要多少就能算多少.在进一步整理了自己的思路之后,他取出一块新的泥板,不仅在上面记录下了15组这种三数组,而且简要地提示了某些计算步骤.但是他没有写出他的新方法的细节,这些要留到给同事们做报告时再说.那时他的同事们就不得不承认他的能力,这样他作为最好的数学教师之一的声誉就会传遍整个王国.

上面开头那句摘自一部古代埃及的珍稀数学文献的引语,以及那个讲叙巴比伦书记员的虚构故事表明了想准确描绘古代数学的某些难处.实际上,凡是有记载的古代文明就一定有数学,但是它总是落在那些受过专门训练的祭司和书记员以及政府官员们的手中,这些人的职务就是在收税、测量、营造、制订日历,祭仪等等领域内来发展和利用数学为政府谋取利益.尽管许多数学概念起源于它们在这些场合的应用,但是数学家们总是要发挥他们的好奇心,把他们的思想推进到超过实际需要的范围,不过由于数学是权力的工具,它的方法只传授给有特权的少数人,而且常常是通过口传.因此文字记载通常是很珍稀的,不大能提供细节内容.

近年来,学者们下了很大的功夫从一切可能找到的蛛丝马迹来重建古代文明中的数学.自然,他们不会在每一点上都意见一致,但已有足够的共识使我们能够对埃及,美索不达米亚,中国和印度的古代文明中的数学勾勒出一幅合情合理的图画.为了能清楚地看出这些古代文明中数学的相

似和差异,我们将不把这些文明分开来论述,而是围绕着下面一些关键性的论题来组织我们的讨论:计数,算术计算,线性方程,初等几何,天文和历法计算,平方根,"毕达哥拉斯"定理和二次方程.为了明了整个故事的背景,我们先来谈谈这些文明本身,谈谈我们从中获知它们的数学的材料来源.

1.1 古代文明

世界上最古老的文明可能就是美索不达米亚文明,它大约在 3500 B.C.出现于底格里斯和幼发拉底的两河流域.在随后的 3000 年中,这个区域先后出现过许多王国,其中有一个建于巴比伦城,它的国王罕谟拉比(Hammurapi)在大约 1700 B.C.左右征服了这整个地区(图 1.1).这个时期的文明已经有了忠于国家的政治观念和众神的思想,这里包括恩里尔(Enlil)神,它是美索不达米亚传统宗教首府尼普尔(Nippur)的市神,还有马尔杜克(Marduk),它是巴比伦城的市神.出现了官僚集团和职业军队,而且在农民大众和王室官员之间一个由商人和工匠组成的中间阶级崛起了.此外,作为会计工具,书写被发明出来了,它随后又被用于协助政府维持对广大地区的集中统治.书写是用一支尖笔在泥板上进行的.在过去的 150 年中共挖出了成千上万件这样的泥板(图 1.2).第一个能够将这种楔形文字翻译出来的是罗林森(Henry Rawlinson,1810—1895),他在 19 世纪 50 年代中期,通过比较分别用波斯文和巴比伦文来描述国王大流士一世一次军事胜利事迹的碑文做到了这一点,碑文刻写在贝希斯敦(Behistun,今伊朗境内)的一块岩壁上.



图 1.1 一张伊拉克邮票上的罕谟拉比(Hammurapi).



图 1.2 一张奥地利邮票上的巴比伦泥板.

有大量这种泥板上刻有数学问题及其解,或专刻有数学表.有几百片已被传抄、翻译并得到阐释.它们的形状通常为长方形,但有时也有圆的.它们拿起来很合手,厚度约为一英寸左右,尽管它们有的小得像张邮票,有的又大得像本百科全书.这种泥板不易毁,这真是一件幸事,因为它是我们了解美索不达米亚数学惟一的来源.这些文字记载的传统在希腊人的统治下于公元前的几个世纪就失传了.这种状况一直持续到了 19 世纪.绝大部分挖掘到的泥板都是来自罕谟拉比时期的.当然也有一小部分来自美索不达米亚文明的最早期,还有来自 1000 B.C.左右几个世纪和在 300 B.C.左右的塞琉古时期的.本章中的讨论多半是有关"古巴比伦"时期(罕谟拉比时期)的,不过,我们使用"巴比伦的"这个形容词指的是美索不达米亚文明和文化.尽管巴比伦只在一段有限的时间内是这个地区的重要城市,这在数学史中已成为一个惯例了.

大约 7000 年以前在埃及的尼罗河流域就出现农业.但是同时统治着**上埃及**(河谷地区)和**下埃及**(尼罗河三角洲地区)的第一个王朝出现于 3100 B.C.左右.第一代法老们遗留下来的是一群由官员和祭司组成的精英,一座富有的宫廷,还有,为这些国王们自身留下了一个介于人与神之间的地位.这种地位促进了埃及不朽的巨型建筑的发展,包括作为王室寝陵而建造的金字塔以及在卢克索和卡纳克建造的大寺庙(图 1.3).那些书记员们逐步创造出象形文字,它们点缀着陵墓和寺庙.商博良



图 1.3 吉瑟地方的金字塔.

(Jeam Champollion,1790—1832) 在 19 世纪初作为主要贡献者首次翻译了这种文字. 他是借助于多语种的碑文 —— 罗赛塔石碑 —— 来完成这项工作的,这些碑文用的是象形文、希腊文,以及稍后出现的通俗文字,一种类似于纸草书上僧侣用的那种简化象形文字(图 1.4).



图 1.4 丁 · 商博良与罗赛塔石碑.

不过我们关于古埃及数学的知识大多数不是来自那些寺庙里的象形文字,而是来自两本纸草书,其中收集有数学问题及其解答:一本是《兰德数学纸草书》,它是苏格兰人兰德(A.H.Rhind,1833—1863) 于 1858 年在卢克索购得的,故以他命名;另一本是《莫斯科数学纸草书》,它是戈罗尼雪夫(V.S.Golenishchev,殁于 1947) 于 1893 年购得,随后赠给了莫斯科艺术精品博物馆. 前者是由僧人阿默士(A'h-mose) 在 1650 B.C.左右从距当时 200 余年前的一部原著上抄录的,它长约 18 英尺,高约 13 英寸. 后一纸草书差不多也属同一时期,长 15 英尺多,高只有 3 英寸多一点. 和美索不达美亚泥板的情况一样,多亏了埃及气候干燥,使得这两本纸草书和其他几百件纸草书得以幸存下来,因为恰恰又是希腊人在公元开始前后的几个世纪里对埃及的统治导致了埃及本土简化象形文字的消亡(图 1.5).



图 1.5 阿门霍特普,一个埃及高官和书记员(公元前 15 世纪).

虽然有传说把中国的文明推溯到五千多年以前,但是这个文明最早的可靠证据是由黄河附近的安阳出土文物所提供的,其年代定在 1600 B.C.左右. “甲骨文” 就属于以该处为中心的商代社会,这是一些奇异的骨片,上面刻有远古时代的文字,它们是那个时期的祭司们用来祭神的. 这些甲骨是我们关于中国计数制知识的来源. 大约在公元前第一个千禧年开始,商朝被周朝取代,它随后又分裂成为许多互相对立的封建国家. 在公元前 6 世纪,出现了一个百家争鸣的时期,其中最有名的哲学家就是孔子. 在好些国家里都建立了学者们的学术团体,在由于铁器的使用而引导的技术成长时期,个别学者还被别的封建诸侯聘用来当他们的谋士.

随着一些弱小的诸侯国逐渐被强国所吞并,这个封建战国时代就结束了,最后到 221 B.C.秦始皇统一了全中国. 在他的领导下,中国转变成了一个高度集中的官僚体制的国家. 他强化了严厉的法制,公平税赋,统一货币和度量衡,特别是统一了文字. 传说认定,秦始皇曾下令焚烧先前传下来的所有的书籍以压制异议,但是我们有一定的理由怀疑这个法令是不是真的实施了. 秦始皇死于 210 B.C.,他的王朝很快就被推翻,并为汉朝所取代,这个汉朝维持了近 400 年. 汉朝建立了由训练有素的公务人员组成的机构,为此需要有相应的教育体系. 在用于培训的教科书中有两本数学著作,可能是在汉朝的早期编就的,一本叫《周髀算经》,一本叫《九章算术》. 要想准确地确定这两本书所包含的数学内容的发现日期是不可能的,但是有一些零星的更古老的资料记载了类似于九章的内容,一般认为至少有些内容周代伊始在中国就已经存在了. 自然,必须牢记,即使按这个判断年代,我们所讨论的中国数学的发展至少比在美索不达米亚和在埃及的发展晚好几百年. 这些文明是否有些东西传到过中国谁也不知道.

公元前 3000 年在印度河两岸有一个叫做哈拉帕的文明兴起了,但没有看到有关于其数学的直接证据. 有这种证据的最早的印度文明是由迟至公元前 2000 年从亚洲草原地带迁徙过来的雅利安部落在恒河两岸形成的. 在大约公元前 8 世纪,君主政体的国家在这个地区建立起来了,而且他们还得管理复杂的系统,例如城堡建设,集中化的行政管理以及大型灌溉工程. 这些国家有等级森严的社会制度,由国王和僧侣(婆罗门) 来统治. 婆罗门的文学多少代以来都是口述,以一种叫做“吠

陀”的长诗的形式来表达. 尽管这些诗歌可能在公元前 600 年就已经取得了眼下的形式,但在纪元以前还没有书写的记载(图 1.6)



图 1.6　一份《吠陀》原稿.

一些吠陀时代的材料描绘了祭司们繁褥的献祭仪式. 就是在这些叫《绳法经》(Sulvasutras) 的著作中,我们发现了数学的概念. 奇怪的是,尽管这里数学是讨论用砖来造祭坛的理论的,但早期的吠陀文明并没有制砖工艺的传统,而哈拉帕文化却有. 因此《绳法经》中的数学有可能是在哈拉帕时期创造的,虽然至今还不清楚它们是怎样传承到这后一时期的. 不管怎么讲,我们有关印度数学知识的来源就是《绳法经》.

尽管在公元前一千年以前世界上其它地方也有文明,但迄今所发现的资料对它们的数学并未提供什么线索. 因此只有等到有了新的考古发现后才能来作进一步的讨论.

1.2 计　数

最简单的数学概念 —— 这可能在有文明以前就已存在 —— 就是计数,用话语来计数. 或以更永久的方式用书写的符号来计数. 虽然研究各种不同语言中的数目字也是很有意思的(见 1.1),我们在此仅限于讨论数的符号. 这些符号的书写有几种不同的组织方法. 一种方法叫编组法. 最简单的编组形式就是用一道斜杠 / 来代表数 1;作特定的重复就可以表示较大的数. 这种数的表示法最早出现在扎伊尔的伊尚果(Ishango) 地方发现的一块骨化石上,用碳测定年代法定出其年代大约在公元前 20 000 年左右. 不清楚这些骨头上的划痕或缺口具体表示什么东西的数目,但是一位仔细研究过这种骨化石的学者认为,它们是对月亮的某些周期进行计数.[3] 早期人用这种初步的数量表示方式来记载天文现象这一点能进一步证实 —— 如我们在晚得多的时期所看到的那样 —— 数学的发展是与天文学的发展携手并进的. 在欧洲的中部也发现了类似于伊尚果骨的人工制品,年代可能在 8000 B.C.,上面有规则分组的刻槽,可能也是表示天文观察结果的.

用编组法表示数的一个更为复杂的例子是埃及人在大约 5000 年前发明的. 在这个象形体系中,十的头几次幂由不同的符号表示. 开始用我们熟悉的一竖表示 1. 然后用∩来表示 10,用𓍢表示 100,用𓆼表示 1000,用𓂭表示 10 000. 通过适当地重复这些符号就可表示任意的整数. 例如,为了表示 12 643 这个数,埃及人就会把它写成 |||∩∩∩∩𓍢𓍢𓍢𓍢𓍢𓍢𓆼𓆼𓂭. (注意通常习惯是把小的数位放在前面.)

象形数系是用来写在寺庙墙壁上或刻在柱子上的. 但是当书记员要在纸草本上书写时,他们就需要一种手写书法,为此他们创造出了一种简化的象形数系,成为符号记数系统的一个例子. 在这种数系中从 1 到 9 的每一个数字都有一个特定的符号,同样从 10 到 90 的每一个 10 的倍数以及从 100 到 900 每一个 100 的倍数也都有各自的特定符号,由此类推,一个给定的数,例如 37,就是在 7 的符号之后写上 30 的符号. 由于 7 的符号是ℓ,而 30 的符号是ƛ,37 就写成ℓƛ. 又如,由于 3 是写成ɯ,40 写成∸,以及 200 写成ノ,于是 243 的符号就是ɯ∸ノ. 虽然在符号数系中不一定要有表示零的符号,埃及人还真有这样一个符号. 不过它并不是出现在数学纸草书中,而是出现在关于建筑的纸草书中,在那里它是用来标记金字塔结构中底部水平线的,这样的符号还出现在记载会计帐目的纸草书中,用来在平衡表中表示支出和收入相等.[4] 在希伯来文和希腊文中使用了类似的符号数系,在这两种语言中所有特定的符号都是直接用字母表中的字母来表示.

补遗 1.1

不同语言中的数目字		
18		
英文	eighteen	8. 10(eight = 8, teen ⇒ ten = 10)
威尔士文	deu naw	2 × 9(deu 来自 dou = 2, naw = 9)
希伯来文	shmona-eser	8. 10(shmona = 8, eser = 10)
约鲁巴文	eeji din logun	20 减 2(ogun = 20, eeji = 2)
中文	shih-pa	10. 8(shih = 10, pa = 8)
梵文	asta-dasa	8. 10(asta = 8, dasa = 10)
玛雅文	uaxac-lahun	8. 10(uaxac = 8, lahun = 10)
拉丁文	duodeviginti	20 去 2(duo = 2, viginti = 20)
希腊文	okto kai deka	8 和 10(okto = 8, deka = 10)
40		
英文	forty	4 × 10(这里 10(ten) 变为 ty)
威尔士文	de-ugeint	2 × 20(de 来自 dau = 2, ugeint = 20)
希伯来文	arba-im	4*s*(arba = 4, im 是复数字尾)
约鲁巴文	ogoji	20 × 2(ogun = 20, eeji = 2)
中文	szu-shih	4 × 10(szu = 4, shih = 10)
梵文	catvarim-sat	4 × 10(eatvarah = 4, sat 来自 dasa = 10)
玛雅文	ca-ikal	2 × 20(ca = 2, kal 是 20 的字尾)
拉丁文	quadraginta	4 × 10(quad = 4, ginta 来自 decem = 10)
希腊文	tettarrakonto	4 × 10(tettara = 4, kunta 来自 deka = 10)

本表列出了在九种古代和现代语言中表示 10 和 40 的词,并对其字源作了语义分析[2].

自有史料记载的时期以来中国人就使用了一种累积式的记数方法,也是以 10 的幂为基础的.就是说,他们创造了一套记数符号来分别表示从 1 到 9 的各个数和 10 的各次幂.于是,例如,数 659 的写法就是:将 6 的符号(个)附加到代表 100 的符号(⊚)之上,然后与 5(区)和 10 的符号(/)放到一起,最后是写代表 9 的符号$:.这种记数系统的发展可能与早先使用的算板有关,在这种算板上算筹排成竖列,代表 10 的各个幂次(见补遗 1.2).于是,很自然地把 659 看成在代表 100 的列中有 6,在代表 10 的列中有 5,而在代表 1 的列中有 9,在书写形式中就各用一个特殊的符号来代表每一列.

这类记数系统的任一个,除了第一个以外,都是围绕着基数 10 来组织的.就是说,对 10 的整数幂都各有不同的符号,其他数的符号就是由这些符号来构成的.巴比伦人则用两种不同的基数为基础来构成他们的记数符号.首先他们采用以十为基础的编组系统来表示小于和等于 59 的数.于是在泥板上垂直的一笔划 Y 代表 1,而一个扭歪的笔划 ≺ 就代表 10.在编组记数法中 37 就表示成 ≺≺≺YYYYYYY.接着在公元前第三个千年间的某个时期,巴比伦人对大于 59 的数创造了第一个定位式或者说位值记数系统.在这个记数系统中,基数 —— 在此为 60—— 的幂不是由符号而是由"位置"来表示,而在每个位置上的数字表示这个位上幂的个数.这样,巴比伦人把 13 329(等于 3×60^2

+ 42 × 60 + 9) 表示成 𒐗 𒐏𒐖 𒐝 (今后我们将把它写成 3,42,05,而不用巴比伦人的笔划来表示). 古代巴比伦人没有 0 的符号,但如果某个给定数中没有某一幂时,就在那里留一个空位. 在一个数的末位之后就没有空位了,因此就很难区分 $3 \times 60 + 42$(3,42) 与 $3 \times 60^2 + 42 \times 60$(3,42,00),不过有时候他们会在这个数的后面加一个适当的数量词来指明它的绝对大小. 例如"3 42 六十"就表示 3,42,而"3 42 三千六百"就表示 3,42,00. 但是,巴比伦人从未用过一个符号来代表"无"这个意义上的零,有如我们写 $42 - 42 = 0$.

补遗 1.2

中国算板

中国算板就是在上面摆了一些算筹(长约 10 cm 的小竹笺)的台板,通过摆弄这些算筹来作各种计算. 用于表示小于 10 的整数的算筹有两种可能的排列:

1	2	3	4	5	6	7	8	9
𝍠	𝍡	𝍢	𝍣	𝍤	𝍥	𝍦	𝍧	𝍨
𝍩	𝍪	𝍫	𝍬	𝍭	𝍮	𝍯	𝍰	𝍱

为了表示大于 10 的数,中国人在算板上使用了一种进位系统. 把算筹排成一列一列,最右面的列表示个位,接下去的列为十位,再接下去的列为百位,以此类推. 在一排的列中,空列表示零. 为了使操作人员易于读数,上述算筹的两种排列方式交替使用. 竖排式用于个位、百位和万位,而横排式用于其余的位. 因此,1156 由 𝍩𝍠𝍭𝍥 来表示,而 6083 由 𝍮 𝍰𝍢 来表示.

算术运算的做法是,把要运算的数在算板上排成不同的行,然后作适当的操作. 例如,将 6 和 9,即 𝍥 和 𝍨 相加,两根水平筹加起来得十,而竖式的筹摆在一起得五. 多位数加与减的运算一般由左向右进行. 对于乘法,作运算的人要记住一些基础的相乘结果,然后计算过程也是从左到右进行,在每次相乘后再相加. 除法也是类似地来做.

负数在算板上的表示办法是,用某种特征来区别"负"筹与"正"筹. 一种办法是用红筹代表正数,用黑筹代表负数. 用筹来计算的办法本章稍后还要讲. 在算板上的运算终于发展成为这样一些方法,我们可以用它们来解线性方程组和求多项式方程的数值解.

巴比伦人把他们的 60 进位记数系统用于作计算. 但在日常用语中他们经常使用的是一个可能更老一些的记数系统,编组的方式略有不同,而且 1,10,60,100,600,1000,3600 等等都各有不同的符号. 这个记数系统可能是从粘土票券上的记号发展过来的,这种土制票券在公元前三千纪时曾经在美索不达米亚地区使用过[5],是用来记载从一处运送到另一处去的货物的数量的. 起初在这些票券上画有表示货物种类的象形文字,例如油罐. 这样五个椭圆就表示五罐油. 但是这个办法逐渐演变成只用一个符号代表货物,而用好几个符号来表示各种大小的数目. 如果该货物要用几种不同的单位来度量才比较方便(譬如谷物就是这种情况),那么对每种单位就用不同的符号.

有些度量单位,当然不是所有的单位,它们之间有这样的关系,即一个"大的"单位等于 60 个"小的"单位. 那么到了一定的阶段这个记录数目的系统就会发展成这样的形式,同一个数字 1 也能代表 60. 我们不知道为什么巴比伦人要用一个大单位表示 60 个小单位,并采用这个办法来构成他们的记数系统. 一种猜想是,60 可以被很多较小的整数除尽. 因此这个"大"单位的分数值可以很容

易地表示成这个“小”单位的整数倍.巴比伦人的60进位制仍然在我们的角度和时间的计量单位中被采用,这种单位在天文学的文献中保留了多少个世纪,而且至今仍然是世界文化的一个不可替代的部分.

古印度的记数系统未见经传,但是有文字迹象表明,表示数的符号的确存在过.只是到了公元前3世纪才有了书写数字的实例.起初这是个混合记数系统.对从1到9和从10到90的各个数各有一个类似于象形文字的符号,对于更大的数就采用类似于中国的累积式记数系统.例如,200的符号就是2和100这两个符号的组合,而70 000的符号就是70和1000这两个符号的组合.在第六章中将会谈到,近代的十进位制是在印度或邻近的地方产生的,但这已经是大约7世纪的事了.

1.3　算术计算

一旦记数系统出现了,所有各种文明就会制定出各种相应的基本算术运算的规则——加法、减法、乘法和除法的运算规则,而且作为最后一种运算的副产品,还会得出分数的书写和运算规则来.这些运算规则可以看成是一些最早的算法.

一个**算法**就是一张设计好了的有序的指令表,目的是为了得出某一类问题的解,古时候的人做出了处理许多不同问题的各种算法.实际上,古代数学在本质上可以看成就是算法,这和希腊的数学不同,后者强调的是理论.在绝大多数可以得到的古代数学文献中,作者首先是描述要解决的问题,然后用一个算法进行计算以得出其解,这个算法可能是显式的,也可能是隐式的.在这些文献中很少谈到这些算法是怎么来的,它为什么有效,或它们有什么限制.相反,我们只看到许许多多应用这些算法的例子,经常情况是越来越复杂.不过在讨论这些算法时,我们会讲述它们的可能来源和它的道理何在,同时还会谈谈,那些巴比伦人,中国人和埃及的书记员们在面对学生们提出“为什么?”这个永恒的追问时将可能做出怎样的回答.

在埃及人的象形编组数系中,加法十分简单:只要把个位数,然后是十位数,再然后是百位数等等组合在一起就是.当某种符号组的数目有十个时,就用下一个符号来代替.因此,把783与275相加,就是把|||∩∩∩∩∩∩∩∩𓍢𓍢𓍢𓍢𓍢𓍢𓍢与|||||∩∩∩∩∩∩∩𓍢𓍢放到一起,得到||||||||∩∩∩∩∩∩∩∩∩∩∩∩∩∩∩𓍢𓍢𓍢𓍢𓍢𓍢𓍢𓍢𓍢.由于有十五个∩,将其中十个换成一个𓍢.这就使得后一种符号有十个,再把它们换成一个𓆼,最后的答案就是||||||||∩∩∩∩∩𓆼,或1058.减法可以类似地来进行.自然,在此情形,当需要“借”位时,被借的一位符号就转化为十个低一位的符号.

在简化象形记数体系中,相加和相减的运算就不可能是这样简单的算法了.数学纸草书中没有为这些运算提供什么线索;相加与相减的结果都是直接写下来的.很可能书记员们有加法表.在某个阶段这种表可能以书面的形式存在过,但是一个有竞争力的书记员自然会把它们记住.这些书记员可能反向使用过这种表来计算减法的问题.

埃及人的相乘算法是以连续加倍的方法为基础的.将两个数 a 与 b 相乘,书记员的做法是,先写数对 $1, b$,然后重复对这对数的每一个进行加倍,直到下一次加倍会使这一对数中的第一个元素超过 a.然后,在确定了哪几个2的幂次相加会得 a 之后,把相应的 b 的倍数相加,就得到了所需要的答案.例如将12与13相乘,他可以先写下下面的各行数:

1	12
2	24
4	48
8	96

写到最后一行他就可能会注意到,下一步的加法会在第一列中得出 16,这是大于 13 的.然后他会查出哪几次加起来会得 13,即 1,4 和 8,他再把另一列中相应的数加起来.结果写下来就是:总计 13 156.

和前面讲的一样,书记员是怎样进行加倍计算的并没有记载.答数就是直接写在下面.很可能书记员已经记忆住大量的 2 的倍数表.事实上,有一些迹象表明,在埃及南部的非洲区域内,加倍是一种标准的计算方法,所以有可能埃及的书记员是从他们的南方同事那里学来的这种方法[6].此外,这些书记员们多少觉察到每一正整数可以惟一地表示成 2 的幂之和.这一事实保证了这一方法的合理性.它是怎样被发现的呢?我们最好的猜想是,它是通过实验发现,然后由口头传承下来的.

因为除法是乘法的逆运算,一个像 156 ÷ 12 的问题就可以表达成“乘 12 以获得 156”.那么书记员就会写下与上面列出的相同各行.不过这次他要检查右边一列中加起来会等于 156 的各个数;在本例中为 12,48 和 96.那么左列中与之对应的数,即 1,4,8,相加得出结果为 13.自然,相除不一定“得出整数”,在得不出整数时,埃及人就用分数.

埃及人用的分数类型是单位分数,或者说是“几分之一”(分子为 1 的分数),惟一的例外是 2/3,可能是因为这种分数是最“自然的”.分数 $1/n$(n 分之一)在象形文字中表示如下:在整数 n 的符号上方放上符号𓂋.在简化象形文字中用一点来代替.这样 1/7 在象形文字中记为$\overset{\frown}{|||||||}$,而在简化象形文字中记为$\dot{\ell}$.唯的一例外,2/3,有一个特殊的符号:𓂚(在象形文字中)或ϒ(在简化象形文字中)(前一个符号是指它是 $1\frac{1}{2}$ 的倒数).不过,在本书的以后各部分将用 $\bar{n}$ 来代表 $1/n$,用$\bar{\bar{3}}$ 来代表 2/3.

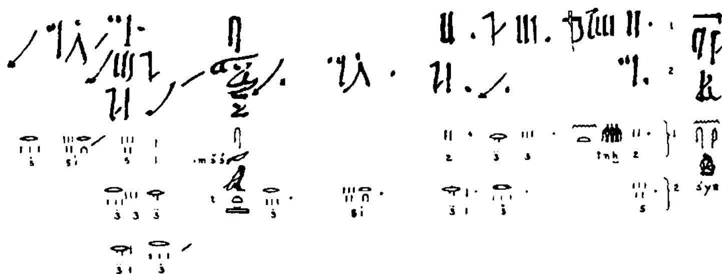
因为分数是作为除不尽时的结果出现,这就要求我们能够处理除单位分数以外的分数.埃及数学最复杂的技巧就是与此相关而发展起来的:即用单位分数来表示任意分数.不过埃及人不是这样来提问题的.当要用到一个非单位分数时,他们就会直接写下许多单位分数之和.例如,《兰德数学纸草书》中的问题 3 是问,如何给 10 个人分配 6 个面包.答案是给每个人 $\bar{2}\ \overline{10}$ 个面包(即$\frac{1}{2}+\frac{1}{10}$).书记员将这个值乘以 10 就可以检验这个结果对不对.我们可能认为这个答案比我们的答案 3/5 要繁琐些,但在某种意义上按这个结果来作实际分割会更容易.如果我们将 5 个面包对半分,再把第六个面包分成十份,然后给每个人半个,再加十分之一个,那么大家都很清楚,每个人都分得了同样多的一份面包.不管繁琐与否,埃及人的这种单位分数之和的办法在整个地中海区域沿用了两千多年.

在将整数相乘时,重要的一步是加倍.同样,书记员也要能够表达出任何单位分数的二倍.例如在上述问题中,对解的验算如下:

1	$\bar{2}\ \overline{10}$
`2	$1\ \bar{5}$
4	$2\ \bar{3}\ \overline{15}$
`8	$4\ \bar{\bar{3}}\ \overline{10}\ \overline{30}$
10	6

这些二倍数是怎样得出来的呢?加倍 $\bar{2}\ \overline{10}$ 简单:因为每一个分母都是偶数,只要将它减半就行了.可

2 DIVIDED BY 3, 5, AND 7



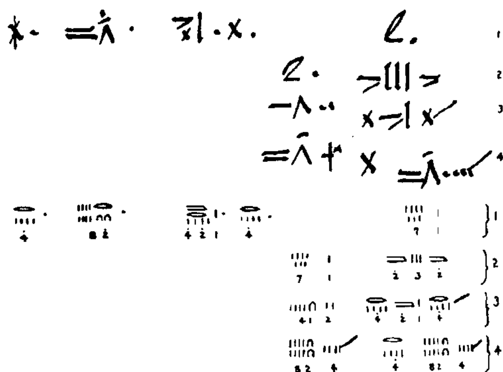


图 1.7　兰德数学纸草书上 2 ÷ 3,2 ÷ 5 和 2 ÷ 7 的摹本及其象形文字译文(来源:兰德纸草书,N.C.T.M).

是,在下行必须将$\bar{5}$加倍.要进行这种计算,书记员就必须用一张表来得到其答案$\bar{3}\,\overline{15}$(2·1/5 = 1/3 + 1/15),事实上,《兰德纸草书》的第一节就是2被从3到101的奇数相除的结果表(图1.7),而埃及的书记员们还认识到了,将$\bar{n}$乘以2的结果就是用n来除2.尽管不知道他们是怎样把这张表构造出来的,已有几种学术上的设想对这种书记员们的'方法'提出了解释[7].不管怎么说,解问题3需要两次使用这张表,第一次上面已经讲过了,第二次,就是接下去这一步,计算$\overline{15}$的二倍,得$\overline{10}\,\overline{30}$

(即2·1/15 = 1/10 + 1/30).这个问题的最后一步涉及把1$\bar{5}$加到4 $\bar{\bar{3}}\,\overline{10}\,\overline{30}$上去,而在此书记员只是给出了答案.人们再次猜想可能有这类相加问题的详尽的表格.有一份埃及数学羊皮卷,其年代大约在1600 B.C.,里面就有这种加法表的一个简化版本.还有几张其他的涉及单位分数的表,和一张处理特殊分数2/3的乘法表.这样看来,埃及书记员们用的算术运算的算法包含着用于加、减和加倍的表格的广泛知识和一套确定的方法,把相乘和相除的问题分解成好几步,而每一步都可以用这些表来完成.在处理相除时,书记员们常常把加倍的程序换成减半的程序.例如,在求2 ÷ 7时,头几步是

$$\begin{array}{ll} 1 & 7 \\ \bar{2} & 3\,\bar{2} \\ \bar{4} & 1\,\bar{2}\,\bar{4} \end{array}$$

为了在右列中得到总和为2,要求在第三行中把$\bar{4}$加到1 $\bar{2}$　$\bar{4}$上去.这样书记员要确定7乘哪个数会得到$\bar{4}$.为此他将已知的结果4 × 7 = 28求倒数,得到7的1/28为1/4.然后,他在他的计算中加上一行$\overline{28}$　$\bar{4}$,并将最后两行相加就得总数为:$\bar{4}\,\overline{28}$　2,即2 ÷ 7 = $\bar{4}\,\overline{28}$.

兰德纸草书上的问题21提供了另一种不同的计算:把$\bar{\bar{3}}\,\overline{15}$补足为1.换言之,我们要确定,在

2/3 + 1/15 上要加上什么数才能得到 1. 书记员注意到 15 的 2/3 是 10, 15 的 1/15 是 1, 加起来总和为 11. 因此他需要一个"乘 15 能得 4" 的数. 步骤如下.

$$
\begin{array}{cc}
1 & 15 \\
\overline{10} & 1\ \bar{2} \\
\bar{5} & 3' \\
\overline{15} & 1' \\
\hline
\bar{5}\ \overline{15} & 4
\end{array}
$$

在这里书记员把第二行加倍就得到第三行. 但是, 由于他认识到 3 的 1/3 是 1, 他取第三行的三分之一就得到第四行. 这样一来, 原问题的解就是: $\bar{5}\ \overline{15}$.

书记员们对他们的基本方法还有一些其他的变更, 可以举兰德纸草书上的题 69 为例, 其中包括了用 $3\ \bar{2}$ 除 80 以及随后的校验:

$$
\begin{array}{cc}
1 & 3\ \bar{2} \\
10 & 35 \\
\text{`}20 & 70 \\
\text{`}2 & 7 \\
\text{`}\bar{\bar{3}} & 2\ \bar{3} \\
\text{`}\overline{21} & \bar{6} \\
\text{`}\bar{7} & \bar{2} \\
\hline
22\ \bar{\bar{3}}\ \bar{7}\ \overline{21} & 80
\end{array}
\qquad
\begin{array}{cc}
\text{`}1 & 2\ 2\ \bar{\bar{3}}\ \bar{7}\ \overline{21} \\
\text{`}2 & 45\ \bar{3}\ \bar{4}\ \overline{14}\ \overline{28}\ \overline{42} \\
\text{`}\bar{2} & 11\ \bar{3}\ \overline{14}\ \overline{42} \\
\hline
3\ \bar{2} & 80
\end{array}
$$

在第二行中书记员利用了他们所采用的记号具有十进小数的性质来求 10 乘 $3\ \bar{2}$ 的乘积. 在第五行中他用了前面提到过的 2/3 乘法表. 于是他发现, 由于第二列中从第三行到第五行之和为 $79\ \bar{3}$, 他还需要在该列中加上 $\bar{2}$ 及 $\bar{6}$ 才能得到 80. 这样一来, 因为 $6 \times 3\ \bar{2} = 21$ 以及 $2 \times 3\ \bar{2} = 7$, 可知有 $\overline{21} \times 3\ \bar{2} = \bar{6}$ 和 $\bar{7} \times 3\ \bar{2} = \bar{2}$, 如第六行和第七行所示. 检验计算表明多次采用了除 2 的表并显示了它在做加法上有很大的方便.

巴比伦人在进行算术计算时用了各种表格, 这一点有大量的直接证据. 许多保存下来的泥板实际上都是乘法表. 不过, 从未发现过加法表. 因为已经分析过的表超过了 200 份, 我们可以认定不存在加法表, 而书记员们是用心记住这些加法步骤并在需要时直接写下答数的. 另一方面, 的确有很多"草稿簿" 的例子, 在上面记载着书记员在解决一个问题的过程中所作的各种计算. 不管怎么讲, 由于巴比伦的数系是一个位值系统, 加、减, 包括进位与借位的实际算法, 很可能与现代的算法很类似. 例如, 将 23,37(= 1417) 加到 41,32(= 2492) 上, 首先是将 37 与 32 相加得 (1,09)(= 69), 这时将 09 写下, 把 1 进到下一列. 同样 23 + 41 + 1 = 1,05(= 65), 从而最终结果为 1,05,09 = (3909).

由于这个位值系统是 60 进位的, 乘法表非常庞大. 这种表的任何一张上就是列出某个具体的数比如说 9 的倍数, 从 1×9 到 20×9, 然后再给出 30×9, 40×9, 和 50×9(图 1.8). 为了计算 34×9, 书记员只要把 $30 \times 9 = 4,30(= 270)$ 和 $4 \times 9(= 36)$ 这两个结果相加, 就可得到结果为 5,06(= 306). 为了将两位或三位的六十进位数相乘, 需要好几张这样的表. 巴比伦人用来作这种相乘的准确的算法 —— 部分积写在什么地方, 最后的结果是怎么取得的 —— 我们都还不知道, 但很可能与我们自己的方法相似.

我们可能以为, 为了得到这种完整的表系, 巴比伦人应对从 2 到 59 的每一个数都造出一张相应

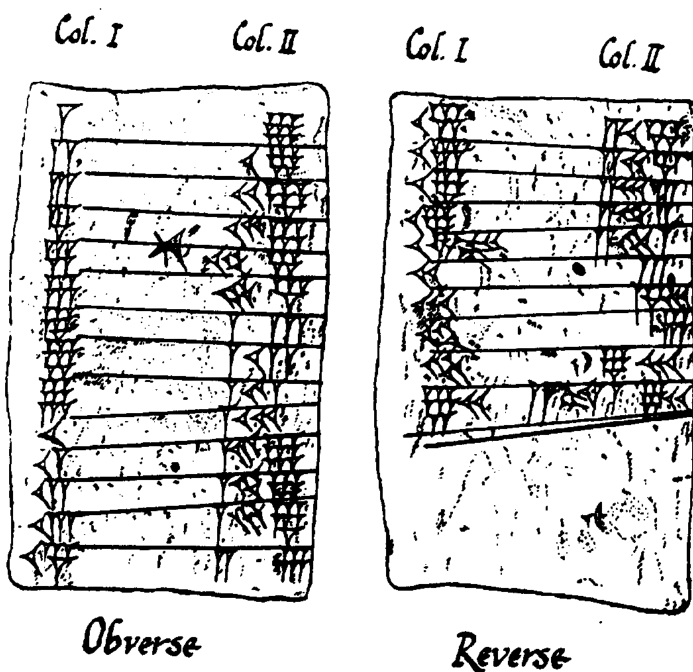


图 1.8 一块巴比伦的乘法表(用于数 9 的)(宾州大学考古系).

的表.然而,事实却并非为此.譬如说,对 11,13 或 17 就没有表,却有 1,15,30,45 的表和 44,26,40 的表.尽管我们不能准确地知道,巴比伦人为什么作了这种选择,可我们却知道,除了 7 这个惟一的例外,至今所发现的乘法表都是对正则六十进位数 —— 即这样的数,它的倒数是一个有尽的六十进分数 —— 而言.巴比伦人把所有的分数都看成是六十进制的分数,类似于我们使用十进制分数.就是说在"六十进制小数点"(记成";")后第一位表示六十分之几,接下去一位为 3600 分之几,等等.于是 48 的倒数的六十进制分数就是 0;1,15,它表示 $1/60+15/60^2$,而 1,21(= 81)的倒数为 0;0,44,26,40,即 $44/60^2+26/60^3+40/60^4$.因为巴比伦人不指明起始 0 或六十进制的小数点的位置,最后的这个数就直接写成 44,26,40.已经讲过,对这些正则数有相应的乘法表.这种表并不表明数的绝对大小,也无此必要.巴比伦人用这种表时自然知道,六十进制小数点的位置有赖于所讨论到的数的绝对大小,所以它的位置最后要根据具体情况来决定.

除了乘法表以外,巴比伦人还使用过大量的倒数表.这里复制了其中一张的部分.一张**倒数表**就是列出一对一对的数,它们的乘积为 1(这里 1 可代表 60 的任意次幂),和乘法表一样,这种表中也只包括六十进制的正则数

2	30	16	3,45	48	1,15
3	20	25	2,24	1,04	56,15
10	6	40	1,30	1,21	44,26,40

曾经将倒数表与乘法表结合起来做除法.因此,1,30(= 90)的乘法表不仅可以用来求该数的倍数,而且,由于 40 是 1,30 的倒数,还可以用它来做除以 40 的除法.换言之,巴比伦人把 $50\div 4$ 的问题看成相当于 $50\times\frac{1}{4}$ 的问题,或者用六十进制的记号,看成相当于 $50\times 0;1,30$ 的问题,下面给出了 1,30 的乘法表的一部分,用它就得出乘积的结果为 1,15(或 1,15,00).再适当地确定六十进制小数

点的位置,就得出这个除法问题的正确答案为 1;15(= 1 1/4).

1	1.30	10	15	30	45
2	3	11	16.30	40	1
3	4.30	12	18	50	1,15

在古代中国,算术计算是用算板来做的.一般来说当需要计算分数时,就要把它们表达成公共分数(common fractions).中国人实际上是使用了我们现代的分数计算法则,包括公分母的办法.然而有一些迹象表明,在早期只是把十进制分数当作算板上的附加列来用的,特别在计算长度和重量时是这样.一个发展得很完善的十进制分数体系还是很久以后的事.

1.4　线性方程

绝大多数的古代数学资料都是有关于问题求解的;为了解这些问题用上了各种各样的数学技巧.作为研究这些问题的开始,我们首先来研究几种用于求解我们今天称之为线性方程的方法.自然,我们必须牢牢记住,古代没有谁用过我们今天用来表示运算或未知量的这样一套符号体系.然而,这些书记员们能够纯粹用文字的技巧来解答这些问题.

埃及纸草书给我们提供了几种处理线性方程的方法.例如,莫斯科纸草书就用了当今的方法来求这样一个数,它的 1 1/2(表示 $1\frac{1}{2}$,译注) 倍再加上 4 就等于 10.用现代的记号,这个方程就是$(1\ 1/2)x + 4 = 10$.书记员的解法和我们今天的作法一样:首先由 10 中减去 4,然后将 6 乘以 2/3(这是 1 1/2 的倒数),得解为 4.类似地,《兰德纸草书》上的问题 31 是要求一个量,它与它的 2/3,它的 1/2 和它的 1/7 相加结果为 33—— 即求 x,使得 $x + (2/3)x + (1/2)x + (1/7)x = 33$.这个问题概念上并不难,但算术运算富有挑战性.这个问题以及接下来的三个问题的目的可能都是为了说明除法方法,因为书记员解这个问题时就是用 $1 + 2/3 + 1/2 + 1/7$ 来除 33.他所给的答案 —— 而这是应加以检验的 —— 为$14\ \overline{4}\ \overline{56}\ \overline{97}\ \overline{194}\ \overline{388}\ \overline{679}\ \overline{776}$(或用现代的记号为 14 28/97).这两个问题是以纯粹抽象的方式提出来的,一点也未涉及到实际的量,例如面积呀,或面包个数呀.事实上,很难找到一个现实生活中的问题的解与这第二个例子有什么关系.书记员只是为了表明他的技术对任何问题,不管有多难,都适用.另一方面,问题35却与实际有关,它需要算出一个斛斗的大小,使用它来量谷物3 1/3次后刚好能填满一赫卡特斗 *[7].书记员所解的这个方程在今天写出来就是$(3\ 1/3)x = 1$,通过将 1 除以 3 1/3 就得解.他把答案 $\overline{5}\ \overline{10}$ 写下并对其正确性进行证明.

解线性方程的第二种技巧在《兰德纸草书》中的问题26中得到了说明,这个问题要求找出这样一个量,它与自身的1/4相加结果为15,解这个问题用的是**假位法** —— 即,先设解为一适当的数,但它不对,然后再对它进行适当的调整.对上述问题书记员的解法如下:“假设[解答为]4.那么 4 加 4 的 $\overline{4}$ 为 5…… 找一个乘 5 能得到 15 的数.答案是 3.用 4 乘 3.答案是 12”[8],用现代的记号,这个问题就是求解 $x + (1/4)x = 15$.第一次猜它是 4,因为 4 的 1/4 是个整数.但是接下来书记员发现 $4 + 1/4 \cdot 4 = 5$.为了求得正确的答案,他必须将 15 被 5 除所得的商,即 3 乘以 4.《兰德纸草书》有好几个这种类似的问题,都是用假位法来求解的.因此书记员所采用的逐步法可以看成是解这类线性方程的算法.尽管没有人讨论这个算法是这样发现的和为什么它能有效,有一点却是显然的,即埃及的书记员懂得两个量之间线性关系这一基本概念 —— 即其中第一个量乘以一个倍数的话,第

* 赫卡特(hekat)是一种谷物干重的度量单位.约等于 1/8 蒲式耳 —— 译注.

二个量也乘以一个相同的倍数.

这一认识在求解比例的问题中得到了进一步的例证.例如,问题75是问:用做155个20比苏的面包同样多的面粉,可以做多少个30比苏的面包?(比苏是埃及人用来度量面包与"强度"相反的单位,可以表示为:比苏 = [面包个数]/[谷物的赫卡特数],其中赫卡特是大约为1/8蒲式耳的干容量.)于是上述问题就成了去计算比例式 $x/30 = 155/20$.书记员是这样来做的:将155除以20,所得结果再乘以30,最后得232 1/2.在《兰德纸草书》和《莫斯科纸草书》的其他地方也有类似的问题.

最后一个来自《兰德纸草书》的线性问题,尽管是特别怪的一个,使用的是一种完全不同的技巧.问题64是这样:"如果对你讲,让你把10赫卡特的大麦分配给十个人,使每个人与他的邻人分到的大麦的赫卡特数之差为1/8,每个人分到的是多少?"[9] 在本题不言而喻的是,分配的数额是按等差级数来分的,这在纸草书中别的类似问题中也是一样的.平均份额是1赫卡特.最大份额可以这样来求:在平均份额上加上1/8乘以份额差的数目的一半.然而,因为这里只有9个份额差,是奇数,书记员的办法是,将平均份额加上公差一半(1/16)的9倍,从而得最大的份额为1 9/16($1\ \overline{2}\ \overline{16}$).由此值分别9次减去1/8,就得到每一个份额,这样完成了问题的全部解.

在中国的《九章算术》第三章中出现过一些商品分配的类似问题.例如,问题1中要求计算在五个官员中按比例5:4:3:2:1分配五只鹿.作者是这样来解这个问题的:把比例中的五个项加起来得15,然后依次将比例中的每一项乘以五只鹿,再被15来除.这样最重要的官员获得 $5 \cdot 5/15 = 1\ 2/3$ 只鹿,第二位官员获得1 1/3只,由此类推.

在《九章》的第二章还有很多类似于埃及人的面包问题的,涉及简单的比例问题.例如,假设50单位的小米可以交换24单位抛光大米,问题3是问:用4 5/10单位的小米能换得多少大米.用现代的记号来写就是要解比例式 $4\ 5/10 : x = 50 : 24$.中国的作者将右边约化为25:12,并如下来解这个问题:12乘4 5/10,然后除以25,和我们的解法一样.

在巴比伦的教科书中很少有单个的线性方程,又没有提供求解的算法.例如,在YBC 4652号泥板上有个问题是:"我找到一块大石,但没有称它;在我给它加上(总重的)七分之一以后又加十一分之一,其重为1米拉(mina)[=60斤(gin)].这块石头原重若干?"[10].我们可以把它翻译成现代的方程 $(x + x/7) + \frac{1}{11}(x + x/7) = 60$.书记员只是直接写下答案,这里为 $x = 48\ 1/8$.可能这类问题解题的过程是记在某些尚未发现的泥板上.

另一方面,对两个未知量的方程组倒是给出了较为详细的解题过程.所用方法中有一个是利用方便的猜测然后再加以调整,这表明巴比伦人也懂得了线性性质.下面是从古老的巴比伦泥板书VAT8389中摘出的一个例子:两块田地中一块每沙尔(sar)出产2/3西拉(sila)谷物,另一块每沙尔产1/2西拉(sar和sila分别为面积和容积的度量).第一块地的产量比第二块的多500西拉;两块地的面积总共为1800沙尔.每块地是多大?以 x 和 y 来代表未知面积,很容易把这个问题翻译成由两个方程组成的方程组:

$$\frac{2}{3}x - \frac{1}{2}y = 500,$$

$$x + y = 1800.$$

一个现代的解法可能是这样来做的,从第二个方程中解出 x,并将其结果代入第一方程,但在这个问题中巴比伦书记员的做法是,开始假设 x 和 y 都等于900.然后计算 $2/3 \cdot 900 - 1/2 \cdot 900 = 150$.所要求的500与150之差是350.为了调整答数这个书记员可能认识到:x 的值每增加一个单位,则 y 的值会减少一个单位,从而"函数" $(2/3)x - (1/2)y$ 就会增加 $2/3 + 1/2 = 7/6$.因此他只需要解方程 $(7/6)s = 350$,得知必须增加 $s = 300$.再900加300得 x 的值为1200,900 - 300得 y 的值为600,

这是正确答案.

中国人也对线性方程组有兴趣,并用了两种基本的方法来处理它们.第一种方法叫盈不足术,见于《九章》的第七章,它主要用于求解用今天的话说就是两个未知量的两个方程组问题.解法和巴比伦的相像,开始也是"猜测"可能的解,然后调整所猜的解以获得正确的解,这表明中国人也已经懂得了线性关系的概念.

现在来讨论问题17:"今有善田一亩价三百,恶田七亩价五百,今并买一顷,价钱一万,问善恶田各几何?"把这个问题翻译成现代的语言就是一组两个未知量的两个方程:

$$x+y=100,$$

$$300x+\frac{500}{7}y=10000.$$

中国人解这个问题的法则是这样说的:"假令善田20亩,恶田80亩,盈$1714\frac{2}{7}$钱;令之善田10亩,恶田90亩,不足$571\frac{3}{7}$钱".这样按照中国作者的说法,其求解过程就是:将20乘以571 3/7,10乘以1714 2/7,将此二乘积相加,最后将此和再除以1714 2/7与571 3/7.这样就很容易求得结果为"善田一十二亩半,恶田八十七亩半".

九章的作者没有说明他是怎样得出这个算法的,这个算法后来首先是在伊斯兰世界中出现,又过了一千多年才在欧洲出现.我们可以将这个算法表示成为下面的公式

$$x=\frac{b_1x_2+b_2x_1}{b_1+b_2},$$

其中b_1是猜测善田为x_1时的盈余数,而b_2为猜测善田为x_2时的不足数.对这一算法是怎样被发现的有一种猜想,认为这可能是基于注意到了未知量的值从正确答案x变为猜想值20时会引起"函数"$300x+(500/7)y$改变1714 2/7,而当从10变到x时又会引起函数值改变571 3/7.由于线性意味着自变量的改变与函数的改变的比值前后两次应该相等,我们得到下述比例式

$$\frac{20-x}{1714\frac{2}{7}}=\frac{x-10}{571\frac{3}{7}},$$

或在一般的情况下,

$$\frac{x_1-x}{b_1}=\frac{x-x_2}{b_2}.$$

由此即推得我们想求的解x.

第7章中的所有20个问题都是用"盈不足术"的这种或那种修正形式来求解的.例如,两个不同的猜测解可能都得出盈余.在每种情况下,作者都对所采用的适合的方法作出了解释.由于所有这种问题用近代的符号来表示都可以写成一种形式,并且有单一的(代数的)解法,读者应该牢牢记住,不论是中国人,还是其他古代的人都从未采用过这种能使类似问题解起来毫不费力的符号系统.所有的问题及其解都是用文字表述的.即使如此,那些书记员们总是迫不及待地推出带有这种笨拙解法的问题,也许是他们想说服他们的学生相信,只要熟练地掌握了这些方法他们就能解决甚至很困难的问题.

《九章》一书的第八章讲述了求解方程组的第二种方法,也是通过各种略有不同的例题来讲解的.不过,在这一情况下,近代方法并不更简单.事实上,中国人的解法实质上与高斯消除法一致,而且是用矩阵的形式表示出来的.例如,考虑该章的问题1:"今有上禾三秉,中禾二秉,下禾一秉,实三十九斗;上禾二秉,中禾三秉,下禾一秉,实三十四斗;上禾一秉,中禾二秉,下禾三秉,实二十六

斗.问上、中、下禾实一秉各几何?" * 这个问题用现代术语翻译出来就是下述方程组:

$$\begin{aligned} 3x + 2y + z &= 29, \\ 2x + 3y + z &= 34, \\ x + 2y + 3z &= 26. \end{aligned}$$

求解的方法是这样讲的:"置上禾三秉,中禾二秉,下禾一秉,实三十九斗,于右方.中、左行列如右方."这个排列用图形表示如下:

$$\begin{array}{ccc} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 1 & 1 \\ 26 & 34 & 39 \end{array}$$

该书继续说:"以右行上禾遍乘中行,而后直除."这就是说,将中间这一列乘以 3(右列中的上禾数),减去右列的一个倍数(在本例中为 2),从而中间列的第一个数化为 0.同样对左列进行类似的操作.结果表示如下:

$$\begin{array}{ccc} 1 & 0 & 3 \\ 2 & 5 & 2 \\ 3 & 1 & 1 \\ 26 & 24 & 39 \end{array} \qquad \begin{array}{ccc} 0 & 0 & 3 \\ 4 & 5 & 2 \\ 8 & 1 & 1 \\ 39 & 24 & 39 \end{array}$$

"然以中行中禾不尽者遍乘左行,而以直除".即用中间这一列对左列进行类似操作.结果为

$$\begin{array}{ccc} 0 & 0 & 3 \\ 0 & 5 & 2 \\ 36 & 1 & 1 \\ 99 & 24 & 39 \end{array}$$

因为上图等价于下述三角形方程组

$$\begin{aligned} 3x + 2y + z &= 39, \\ 5y + z &= 24, \\ 36z &= 99. \end{aligned}$$

作者接着解释了如何从 $z = 99/36 = 2\ 3/4$ 开始,用今天所谓的"回代"法求解这个方程组[12].

遗憾的是,正如在所有原始材料中一样,对为什么这个算法好使,或者说,这个算法是怎么推导出来的,没有说明.我们只能猜想,中国人发现了,将一方程的倍数减去另一方程所得出的新方程组与原来的那个方程组有相同的解.有一点是清楚的,这个方法相当于在算板上用不同方框中的算筹来求解.有人会问,如果这种矩阵运算在某个方框中得出了一个负值,那会怎么样呢?观察同一章中的问题 3 就会发现,这没有关系.对下述方程

$$\begin{aligned} 2x + y \quad\ &= 1, \\ 3y + z &= 1, \\ x \quad\ + 4z &= 1. \end{aligned}$$

这个方法自始至终都完全是可行的.事实上,作者给出了包括正、负数的加法与减法法则."对于减法 —— 如果两数符号相同,则从一数中减掉另一数;如符号不同,则将两者相加;从无中拿走正数就得负数,如从无中去掉负数,就得正数.对于加法 —— 如两数符号不同,则由一数减去另一数;如

* 这里"禾"是指庄稼,"秉"是"捆","实"是"粮食"—— 译注.

符号相同,则二者相加;正数与无相加为正数,负数与无相加为负数”.[13]

作为一个难点不同的例子,考虑最后一个问题 13. 这是一个有六个未知量的方程组:

$$\begin{aligned} 2x + y \qquad\qquad &= s, \\ 3y + z \qquad\quad &= s, \\ 4z + u \qquad &= s, \\ 5u + v &= s, \\ x \qquad\qquad + 6v &= s. \end{aligned}$$

用矩阵方法最后将得到方程 $v = 76s/721$. 如果 $s = 721$, 则 $v = 76$. 只给出了这一个答案. 遗憾的是,不知道中国人是否考虑了取其他值的可能性,或者说是否考虑了有无限多个解的可能性. 不过,在大多数情况下,巴比伦人和中国人都只讨论了未知数与方程个数相等的情况. 至于为什么在上述情况下只产生惟一的解,或者在其他情况下会怎样,这方面的讨论没有任何记载.

1.5　初等几何

记载表明,我们讨论过的所有古代人都知道如何计算简单的直线图形的面积. 书本中有大量用标准公式 $A = bh$ 和 $A = (1/2)bh$ 来分别计算矩形和三角形面积的例子,只是在这里指定长度为 b 和 h 的直线是否互相垂直这一点不总是很明确. 巴比伦人提出了很多公式来计算各种图形的面积,其中包括有相当于今天所谓的系数表 —— 一种反映不同几何图形的某些数学关系的常数表. 譬如,对三角形给出的系数 0;30(= 1/2) 是指三角形的面积是高和长的乘积的一半. 类似地对三角形的高给出系数 0;52,30(= 7/8) 就是指等边三角形的高是底的 7/8,当然这个结果只是近似正确. 正确的倍数应该是 $\sqrt{3}/2 = 0.866$. 事实上,古代的书记员们对长度和面积求得一个合理的近似值就满足了.

估算圆的周长与面积是使用合理近似的一个重要例子. 没有一个简便的方法来估算一个给定直径的圆的周长,或者准确地来计算它的面积. 现代的公式 $C = \pi d$ 和 $A = \pi r^2$ 二者均含有同一常数 π. 显然,周长与半径成正比,面积与半径的平方成正比,但这两个比例常数相同(今天写成 π) 这一点就不那么明显了. 古代人知道周长与直径成正比,有各种书籍表明这点. 至于谈到面积,埃及人是独立于周长来计算面积的. 然而,巴比伦人和中国人是知道一个圆的面积、周长和直径这三者之间的关系的.

在许多巴比伦人的泥板上,圆的周长是当作直径的三倍来计算的. 类似地,在《九章》第一章问题 31 中有“今有圆田,周三十步,径十步” 的说法. 而在希伯来圣经中,“诸王本纪 7:23 谈到公元前 950 年左右所罗门王在位的时期,其中写道:“他铸造了一个大盆,从边缘到边缘为 10 肘尺,盆面呈圆形 …… 一根长为 30 肘尺的线可绕它一周”. 在这里周长也取为直径的三倍. 在某种意义上来讲,这个结果很难理解,因为甚至很粗糙的测量就会表明周长会比直径的三倍大. 这一数值使用了这么长的时间,可能是由于这样算起来方便,而对于实用来讲,它得出的结果精度也足够了. 也许还因为它是一代一代这样传下来的,传统很难改变.

然而,在讨论圆的多数问题中,要求的不是周长,而是面积. 这方面有各种不同的结果.《兰德纸草书》的问题 50 是这样讲的:“直径为 9 的圆田的例子. 面积是多少?减去直径的 1/9,剩下为 8,将 8 乘以 8,得 64. 于是面积为 64”.[14] 换言之,埃及书记员用的公式为 $A = (d - d/9)^2 = [(8/9)d]^2$. 将它与公式 $A = (\pi/4)d^2$ 相比,表明埃及人所取 π 的数值在计算面积时为 $256/81 = 3.16049\cdots$. 埃及人是从哪儿得来这个值的?为什么答案是用 $(8/9)d$ 的平方形式而不是用现代的、以直径的平方乘以一个倍数(在此为 64/81) 的形式来给出?

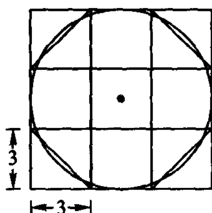


图 1.9 正方形的内接正八边形,采自兰德纸草书问题 48.

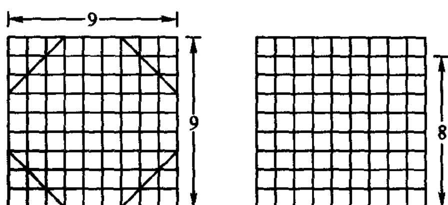


图 1.10 一正八边形的剖分.

还是这本纸草书中的问题 48 给了我们一点启示,在该题中有一个与边长为 9 的正方形内接的八边形(图 1.9).然而,没有陈述问题,光有计算 $8 \times 8 = 64$ 和 $9 \times 9 = 81$.如果书记员在同一正方形中作一个内接圆,他可能发现它的面积与该正八边形的面积接近.由于正八边形的面积是正方形面积的 7/9,可能书记员直接令 $A = (7/9)d^2[= (63/81)d^2]$.显然书记员是想把结果表述成一个正方形面积的形式.他关心化圆为方的问题,即求一个正方形,使它的面积和所给圆的面积相等.因此,他要求这样一个正方形的边,它的面积等于 $(7/9) \cdot 81 = 63$.求这条边长的办法之一是,在边长为 9 的正方形内画一个内接正八边形,再划 81 个小正方形(图 1.10).顶上画了阴影的两个角块与顶行中两个小方块相等,而底部两个小方块等于左列中的两个小方块.移动该行与列(将一方块移动两次),留下来的正方形的边长为原来正方形边长的 8/9,这与正八边形的面积非常接近,因而也就与圆的面积接近.这一重组的图形可能澄清了为什么书记员要坚持"减去直径的 1/9",然后将剩下的加以平方.应当指出,问题 50 不是一个孤立的问题.在其余两个问题中,其中圆只是一个更复杂图形的一部分,也用了相同的方法.

在印度的《绳法经》中,也解决了化圆为方的问题,在这里问题涉及建造一个方形祭坛,使它的面积与一圆的面积相等:"如果你想把一个圆变成一个正方形,将直径分成八份,再将其中的每一份又分成 29 份,除去其中 28 份,再去掉这剩下的一部分的六分之一减去这六分之一的八分之一"[15].印度祭司的这句话的意思是,这要求的正方形的边长为圆直径的

$$\frac{7}{8} + \frac{1}{8 \times 29} - \frac{1}{8 \times 29 \times 6} + \frac{1}{8 \times 29 \times 6 \times 8},$$

这相当于取 π 为 3.0883.

巴比伦人与中国人处理面积问题的方法不同.在这两个文明中,圆的面积是用公式 $A = Cd/4 = (C/2)(d/2)$ 来计算的,其中 C 为周长.他们也都用了公式 $A = C^2/12$,只要用 $C/3$ 代替 d 就很容易导出这个公式.实际上,巴比伦人对圆用的系数是 1/12,圆的面积就是用周长的平方乘这个系数来计算.中国人有时用公式 $A = 3d^2/4$,这可从在公式 $A = Cd/4$ 中用 $3d$ 代替 C 来导出,但也可以用圆的内接正方形和外接正方形的面积的平均值来得出.

巴比伦人和中国人是怎么发现公式 $A = (C/2)(d/2)$,因而将面积的计算与周长的计算联系起来的呢?和经常出现的情况一样,这个问题在教材中没有答案.一种可能的解释是,他们把圆切成扇形,然后把它们重新排列成一个近似的矩阵(图 1.11).另一个可能的解释是,通过将圆分割成许多无限细的同心圆来说明,这有

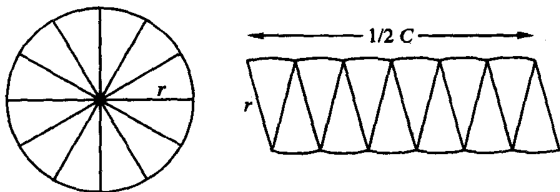


图 1.11 圆的剖分.

中世纪的一些考证作依据.如果我们将圆从圆心开始直到边缘分成一些细窄的圆,并将这些同心圆的狭条展开叠成一个三角形,这个三角形的底为圆周长,高为这个圆的半径.面积的公式立即可以导得(图 1.12).

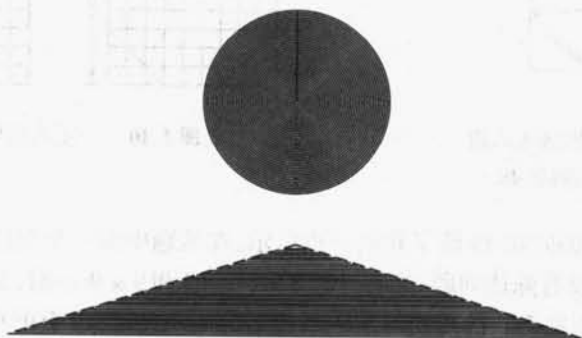


图 1.12　由无穷小的细条构成的圆.

处理平面图形的最后几个结果涉及计算弓形面积的公式.在这一情形,《九章》给出的法则是 $A = (sp + p^2)/2$,其中 s 为弦长,p 为"矢"长.* 这个公式只有在弓形为半圆这个特殊情形下才是对的(假定取 π 为 3).但这个公式用来计算高为 p,上底和下底分别为 p 和 s 的梯形面积倒是正确的.因此这个公式可能是出于用梯形来近似计算弓形面积(图 1.13).有趣的是在公元前 3 世纪的埃及纸草书上也出现过这个公式,后人就把叫做"古法".巴比伦人计算了两种不同的双弓形 —— 一种是"船形",由两段四分之一圆弧围成,另一种叫"牛眼",由两段三分之一圆弧围成(图 1.14),但计算方法大异其趣.这两种图形的面积分别由公式 $(2/9)a^2$ 及 $(a/32)a^2$ 给出,其中 a 为弧长.在假设圆面积为 $C^2/12$ 及 $\sqrt{3} = 7/4$ 的情况下,以上两个式子都是精确的.

在立体几何的公式方面,也有相当广泛的知识.计算长方体的公式 $V = lwh$ 与计算柱体体积法则一样众所周知.事实上,在《兰德纸草书》中的好几个问题里书记员用了计算柱体的体积公式 $V = Bh$,式中 B 为柱体的底面积,它由已经讨论过了的圆面积公式来计算.

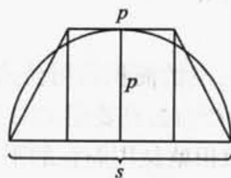


图 1.13　对弓形面积的近似.

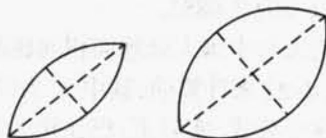


图 1.14　巴比伦船和牛眼.

在巴比伦泥板文书和《九章》中都涉及了墙与坝的体积计算以及建造它们需用多少工人等问题.在很多情况中墙的截面为梯形;这时体积就可由墙的长度与梯形面积相乘来计算.

因为埃及最著名的建筑之一就是金字塔,人们可能希望能找到计算它的体积的公式.遗憾的是,并没有找到这种公式.然而,在《莫斯科纸草书》中倒是有一个与金字塔有关的、引人注目的公式,这就是计算截棱锥体体积的公式:"设一截棱锥体高为6肘,底边长为4肘,顶边长为2肘;先对4

* "矢"长是指弓形的高 —— 译注.

这样计算,将其平方,得16,再将此4加倍,得8.然后来计算2,将其平方,得4,将上述16,8和4加起来,得28.计算6的1/3,得2.再将28二倍,结果为56.啰!它就是56.你已经找到了正确的结果".[16]如果将此算法译成公式,以 a 表下底长,以 b 表上顶长,以 h 表高,则它的结果就是正确的公式 $V=\frac{h}{3}(a^2+ab+b^2)$. 尽管没有哪种纸草书给出过底为正方形的完整的金字塔公式 $V=\frac{1}{3}a^2h$,这里 a 为底边长,h 为塔高,但只要在截棱锥的公式中令 $b=0$ 就立即得出了这个公式.因此我们假定埃及人是认识到了这个公式的.从另一方面来讲,要想从完整的棱锥体的体积公式导出截棱锥体的体积公式,就需要较高的代数技巧.尽管提出了许多涉及剖分的中肯的建议,至今还无法肯定地说埃及人是怎样发现他们的公式的.

《九章》提出了与《莫斯科纸草书》中一样的公式,同时还有完整的棱锥体的公式.3世纪的一位《九章》的评注者用一种聪明的立体剖分方法给出了前一公式的一个证明,但在他的论证中必须用到棱截体体积公式.自然,那位推出这个棱锥体体积公式的中国数学家又是如何证明它的,仍是无人知晓*.

巴比伦人也考虑过立体的体积,包括棱锥体.最好的例子来自泥板 BM 96594,在它上面有几个涉及长方形棱锥状谷物堆,这里棱锥的顶是伸长了的,好像一个斜屋顶(图1.15a),求解的方法相当于下述公式

$$V=\frac{hw}{3}\left(l+\frac{t}{2}\right).$$

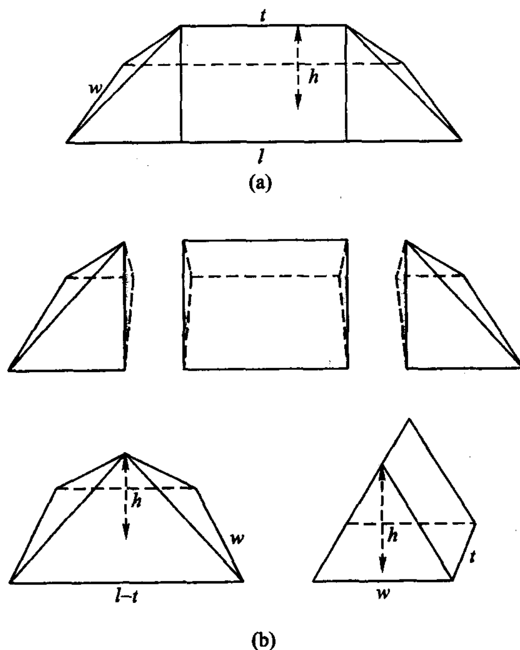


图1.15　巴比伦人的谷物堆及其剖分.

---

* 这里所说的《九章》评注者,应指刘徽.对刘徽关于棱锥体体积公式的推导,D.B.Wagner已给出了一个令人信服的解释,参见D.B.Wagner, An Early Chinese Derivation of the Volume of a Pywnid: Liu Hui, 3rd Century A.D. Historia Mathematica, 6(1979).——译注.

其中 l 是该立体的长，w 是它的宽，h 是它的高，t 为它的顶部的长度. 尽管在泥板上没有这个正确公式的推导，我们能够通过把这个立体剖分成一三角棱柱，两边各加上一矩形棱锥的一半. 于是其体积就是这两个立体的体积之和(图 1.15(b)). 这样一来，V = 三角形棱柱的体积 + 矩形棱锥的体积，或

$$V = \frac{hwt}{2} + \frac{hw(l-t)}{3} = \frac{hwl}{3} + \frac{hwt}{6} = \frac{hw}{3}\left(l + \frac{t}{2}\right),$$

这正是我们想求的.[17] 至今尚未发现上面记有完整的棱锥体积公式的巴比伦泥板，但是和埃及人的情况一样，从我们已讨论过的情况来看，可以合理地假定，巴比伦人是知道其正确公式的.

因为有一块泥板上面有计算一个正方形底面积为 a^2、正方形顶面积为 b^2、高为 h 的截棱锥体积的正确公式

$$V = \left[\left(\frac{a+b}{2}\right)^2 + \frac{1}{3}\left(\frac{a-b}{2}\right)^2\right]\cdot h,$$

上面的假设就更令人信服了. 当然，完整的棱锥公式可以从上式中令 $b = 0$ 推出. 可是另一方面，有这样的泥板，其中计算这种体积的公式是 $V = \frac{1}{2}(a^2 + b^2)h$，这是梯形面积公式的一个简单的，却是错误的推广. 然而，应当记住，尽管这个公式是不对的，但是它算出来的答案和正确的答案相差不大. 由于当时没有精确的体积测量方法，因此很难有人能发觉这个结果是错误的. 无论如何，由于要用到这个公式的问题都是实际问题，常常与要造一特定的建筑所需要的用工数有关，公式带来的小小误差对最终答案并没有多大的影响.

1.6　天文计算

体积问题对埃及人和巴比伦人二者都很重要，这是因为这些问题在建造金字塔、庙宇和防洪设施上有实际的应用. 它对建筑师和工程师们确定这些建筑所需的材料数量来说是不可缺少的，他们可以由此来算出所需的工人数以及供养这些工人所需的面包数. 实际上，古巴比伦的系数表中就包含了用于求解这类问题的一些标准比率，这些比率涉及工程量和应付各类工人的工资数.

这种庞大的土建工程项目很多都是为了祭神的目的而建造的. 这种巨大的宗教建筑遍布全球. 构筑这种巨型建筑要求有相当的技术水平和组织能力. 但是在解决一般的工程问题之前，建造者首先面临选址的问题. 许多巨型建筑都与一些重大的天文事件密切相关，所以我们可以认为建筑师们很熟悉天文学的基础. 这方面的知识不仅在建设大型建筑时要用到，对制订历法也是不可或缺的. 因而天文学的问题在某些数学工具的发展上有决定性的意义，这在后来的历史中也是如此.

古代人对上天知道些什么呢? 当时最重要的天体就是太阳和月亮，显然这二者都是从东方升起，在西方落下，但每一个的实际运动却要复杂得多. 例如，在春分时太阳准确地从东方升起，在夏季却在东偏北升起，到了秋分这一天又从正东升起，可是到了冬季，升起的地方又转到东偏南，以上情况至少在北半球是如此. 在任何地方都会观察到太阳的这一循环是按一定的时间间隔重复的. 只要有这种计算的记录，这一时间间隔，即年的长度，均被确定为约 365 天.

如果你想确定年历中的重要日子，你必须要会观察太阳的位置. 部分地正是由于这个原因，在公元前三千年开始的时候在英国的斯通亨格(Stonehenge) 建造了一座巨大的石庙(图 1.16). 在英国的其它地方以及北欧的一些地方都建有许多类似但较小一些的建筑. 尽管建造这些建筑的理由还不完全清楚，但是大多数的学者都认为，它有一个目的就是为了确定太阳上升和下落的最北和最

南的位置.[18] 例如,公元前 3200 年左右,在爱尔兰,米思郡的纽格兰奇(Newgrange) 建造的一座带廊道的墓宅是这样设计的,使得在冬至前后三到四天的时间里 —— 也只有在这几天 —— 初升的太阳光线能够穿过屋顶的一条缝隙照亮这个建筑的后部(图 1.17). 在一些其他的建筑中,石头之间的连线或一块石头与地平线上的某个重要自然标志物的连线准确标记了冬至日太阳升起或下降的方向.



图 1.16 一张英国邮票上的斯通亨格石庙,表明它是用于天文观测的.



(a)



(b)

图 1.17 (a) 法国卡那克(Carnac) 地方石阵的排列;(b) 爱尔兰邮票上纽格兰奇廊式墓穴在冬至被阳光照射的情况.

从理论上讲,以太阳升起的位置为基础,历书就可以制定出来. 但在多数有记载的文明中,是月球的运动确定了一年内的重要时段 —— 月份. 和太阳一样,月球在东方地平线上升起的位置也是变化的,多年耐心的观察使斯通亨格石庙的建造者们显然能够标记出月球升起的最北和最南的位置,他后来也可能已经注意到了月球升起的最北和最南位置有一个 18.6 年的周期变化,用这个周期可以帮助我们来预测月食. 月食和日食对古代人民来讲意义重大. 对这样惊人的现象有预测的能力 —— 并通过适当祭祀活动让那些被"吃掉" 了天体重现 —— 是祭司们的一项重要职能.[19]

然而,月球在空中表现出来的最突出的特征不是它升起的地点,而是它的月相. 所有早期文明都注意到了月球从小小的一弯新月开始,变成满月,再变成看不见,最后又重新变成一弯新月这个过程所经历的时间,这些观察很可能就是迄今所发现的最早的数字记号的基础. 埃及人与巴比伦人都用月相来确定一年中的月份,但方法不同. 从太阳刚下山开始在西天出现一弯新月,经过各个月相到下一次新月出现所历时间很容易确定,大约为 $29\frac{1}{2}$ 天,遗憾的是,没有 $29\frac{1}{2}$ 的某一整数倍能等于一个太阳年的天数 365,所以要编制一种历书,它能协调月相和由太阳所确定的季节,这一工作并不简单. 早期的埃及人对此作了彻底的简化. 他们采用了一种 12 个月的历书,每个月由 30 天组成,再将 5 天附在末尾以构成 365 天的一年. 这种历书不得不忽略月球的周期变化. 此外,由于一年实际上有 365 1/4 天,最后甚至每年的年历也不能与季节同步. 换言之,从一年的起始将在 1460 年 (4×365) 中经历一个完整的季节周期,这一点埃及的祭司也早就知道得很清楚了. 于是,祭司们为了各种宗教目的,始终保持记录阴历月. 在埃及尼罗河,一年一度的泛滥会在农田中带来大量的泥沙,是农业上最重要的事件,祭司们还发现,洪水总是在黎明前不久一段黑暗时期当东方天空上出现一颗亮星 —— 天狼星之后开始泛滥. 因此他们对此作出精确的预言,而这有助于确立他们的权威.

在巴比伦历书的情况则不一样. 那里的祭司们要想使历书能同时适应太阳和月亮的运动,使得一给定的农事能在每一年的同一月份里出现. 因此月份的天数一般来说是在 29 和 30 天之间交替变化,一个新的月份总是随着黄昏时新月初现的一刻而开始. 因为 12 个这样的月只有 354 天,所以他们决定每过几年增加一个额外的月份. 在早期,他们只是在感到必要时通过临时法令来实施这一

点，但到了公元前8世纪中叶，巴比伦人制定了一个新的历法，规定每19年中有7个闰年，每个闰年由13个月组成．月份的长度有时也要加以调整，以使每19年的235个月的周期中包含有6940天．事实上，巴比伦人知道，月球周期长度的平均值约等于29.53天，而这又等于6940/235，当今犹太历法保留了巴比伦历法的主要内容，只作了一些小的修正，以便与犹太人的法律保持一致．

中国人也知道这个19年的周期，以及其它更长的周期．例如，19个太阳年实际上包含6939 3/4天，他们就用了76年（27 759天）的周期，以使每月的天数为整数．中国人还考虑了月食和行星的周期，并得出了这样的结论：经过31 920年后，“一切事物都结束了，又重新回到了它们的原始状态”．[20]

1.7 平方根

我们来回顾一下埃及人在化圆为方的问题中计算一个平方根的例子．其意思实际是要求一个面积为63/81的正方形．对这个问题的一种看法是，把这正方形改写为一个 $7/9 \times 9/9$ 的矩形．为了把这个矩形变成方形，做法是切去一个边长为7/9的正方形，再把剩下的矩形（$7/9 \times 2/9$）分成两半，将其中之一沿着方形一边放置（图1.18）．这样就得出一个**磬折形**（gnomon），即一个正方形折去一个角．埃及的书记员显然满足于假定边长为8/9的较大的正方形是这个磬折形的足够好的近似．毕竟他不会对 $\sqrt{63/18}$ 本身感兴趣，而只是关心如何用它来作圆面积的近似值．由于他并不知道这个面积近似究竟有多大的误差，他略去这块面积为1/81的小正方形就关系不大了．

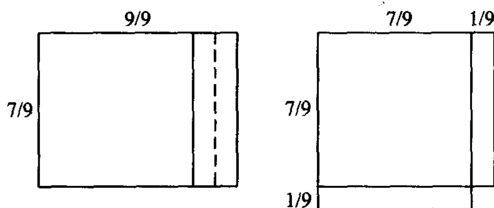


图1.18 将矩形变成正方形．

遗憾的是，没有其他例子表明书记员们真正算过平方根．在需要平方根时，总是把问题改动一下，使得正好能得出平方根来．这不是说，所有这样得出来的平方根都是整数．在一份纸草书中，书记员写下了6 1/4的平方根是2 1/2．很可能是这些书记员们有一张平方表，而这只不过将一张平方表反过来查，这种表是不难制作出来的．如果这张表足够大，书记员甚至能够通过插值求出那些没有在表中直接写出来的平方根，虽然，还不知道有没有这样的例子．

巴比伦人也有广泛的平方和平方根的表，并且对立方和立方根也有类似的表，现在仍存有这样的例子．通常，当在求解问题中需要平方根时，就将问题处理一下，使得要求的平方根在表中，并且是一个有理数．然而，会有这样的情况出现，这时需要的平方根是一个无理数，特别是 $\sqrt{2}$．当出现这个特殊的值时，结果一般都是写成1;25（= 1 5/12）．然而，有一块有趣的泥板，YBC7289，上面画了一个边长为30的正方形，还在它的对角线上写了两个数1;24,51,10和42;25,35（图1.19）．30乘以1;24;51,10所得之积恰好就是42;25,35．这样，可以合理地假设，最后那个数表示对角线的长度，而另一个数就表示 $\sqrt{2}$．

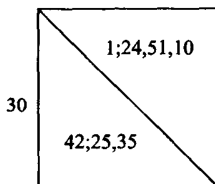


图1.19 巴比伦泥板上的 $\sqrt{2}$．

$\sqrt{2}$是用1;25还是用1;24,51,10来给出,这个数又是怎么算得的,都无据可查.但是由于书记员肯定知道,这两个数的平方都不准确地等于2,也知道它们并不准确地等于面积为2的正方形的边长,他们肯定知道这两个值都只是近似值.这两个值是怎样定出来的呢?一个可能的方法,这个方法在泥板文书上可以查到它的踪迹,它一开始用了一个代数恒等式$(x+y)^2=x^2+2xy+y^2$,这个等式可能是巴比伦人从一个等价的几何结论推出的.现在假设给定了一个面积为N的正方形,我们要来求它的边长$\sqrt{N}$,第一步就是选一个接近但小于要求的值的正则数a.令$b=N-a^2$.下一步就是选择c,使得$2ac+c^2$尽可能接近b(图1.20).如果a^2"足够接近"N,则c^2就会很小(与$2ac$相比),所以可以选c等于$(1/2)b(1/a)$,即,$\sqrt{N}=\sqrt{a^2+b}\approx a+(1/2)b(1/a)$,(为了保持与巴比伦方法的一致,$c$的值写成了乘积的形式,而不是商的形式,由于有一个因子是a的倒数,由此可见a为什么必须是一个正则数.)类似的论证有$\sqrt{a^2-b}\approx a-(1/2)b(1/a)$.在$\sqrt{2}$这个特殊情况下,开始令$a=1;20(=4/3)$.则$a^2=1;46.40$,$b=0;13,20$,而$1/a=0;45$,故有$\sqrt{2}=\sqrt{1;46,40+0;13,20}\approx 1;20+(0;30)(0;13,20)(0;45)=1;20+0;05=1;25$(或17/12).自然,人们实际上并不需要得出$\sqrt{2}\approx 1;25$的具体过程.人们只需要有一个好的估计值,或一张表,其中1;25是一个要作平方的数.总之,$(1;25)^2$与2之差仅为$0;0,25=1/144$.

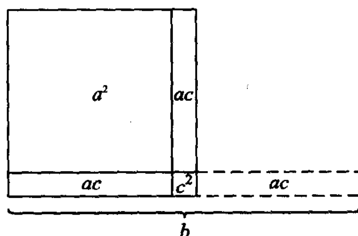


图1.20 $\sqrt{N}=\sqrt{a^2+b}\approx a+\frac{1}{2}\cdot b\cdot\frac{1}{a}$的几何图像.

非常有趣的是,印度的《绳法经》用下式来逼近$\sqrt{2}$:

$$1+\frac{1}{3}+\frac{1}{3\times 4}-\frac{1}{3\times 4\times 34}=\frac{17}{12}-\frac{1}{12\times 34}.$$

这很容易从公式$\sqrt{a^2-b}$并令$a=17/12$来求得,因为

$$\sqrt{2}=\sqrt{\left(\frac{17}{12}\right)^2-\frac{1}{144}}\approx\frac{17}{12}-\frac{1/144}{34/12}=\frac{17}{12}-\frac{1}{12\times 34}.$$

印度人是否用过这个公式无人知晓.不管怎么说,这个值与建造一个正方祭坛有关,该正方体是一给定正方体的二倍.

还不清楚巴比伦人是否也用这个方法计算了$\sqrt{2}$的更好近似值1;24,51,10,因为1;25不是一个六十进制的正则数.有可能他们已经能求得1;25的倒数的近似值,譬如说0;42,21,10,然后计算$\sqrt{2}$如下:

$$\sqrt{2}=\sqrt{1;25^2-0;00,25}\approx 1,25-0;30\times 0;00,25\times 0;42,21,10$$
$$=1;24,51,10,35,25.$$

因为近似公式得出的值略高于真值,书记员们有可能将上述结果切尾后得到1;24;51,10.然而,没有直接的迹象表明他们作了这样的计算,甚至也没有任何迹象表明他们进一步地采用了这种近似算法来计算平方根.

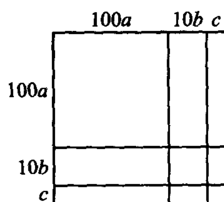


图1.21 中国人的平方根算法.

另一方面在中国,有直接的迹象表明有明确详尽的计算平方根的算法.它也是以代数式$(x+y)^2=x^2+2xy+y^2$为基础.在《九章》的第四章中用文字描述了这个方法.但是研究中国数学史的专家认为,最初的作者可能用了如图1.21所示的图形.[21] 为了解释中国人的算法,我们用问题12来作例子,在该问题中计算了$\sqrt{55\,225}$.求解的思路是这样的,求三个数字a,b,c,使得

结果可以写成 $100a+10b+c$(请记住中国人是用了十进制的).首先,求最高位的数字 a,使得有 $(100a)^2<55\,225$,在本例中 $a=2$,在图1.21中,大的正方形(55 225)与小的正方形 $100a$(40 000)之差为图1.21中那片大的磬折形.如将外层最窄的磬折形略去就可见,b 应满足 $55\,225-40\,000>2(100a)(10b)$,或者说应有 $15\,225>4000b$.故必有 $b<4$.为了校验 $b=3$ 为所求,就是说把边为 $10b$ 的正方形补上所得的那个较大的磬折形面积仍小于15 225,就必须校验是否有 $2(100a)(10b)+(10b)^2<15\,225$.因为这实际上成立,所以可以重复相同的做法来求 c:$55\,225-40\,000-30(2\times200+30)>2\times230c$,也即 $2325>460c$.显然有 $c<6$.简单的校验表明 $c=5$ 给出正确的平方根:$\sqrt{55\,225}=235$.

中国人计算平方根的算法与近代中、小学中所教的方法相似.这个方法给出一系列的答案,在本题中为200,230,235.每一个比它前面一个更逼近真正的结果.所以这是一个确定一收敛数列的例子,数列中的每一个数由一明确的算法从它前面的数来算出.尽管对一个现代的读者来说,如果答数不是整数,则这个过程显然可以用十进分数无限地继续算下去,但当时的中国作者在平方根不是整数时,就用普通分数来作余数.

在此再提出两点评注是适宜的.第一点,仔细考察上述算法表明,这个算法的一部分是一个求解二次方程(或至少是一个二次不等式)的过程.第二,《九章》还有一个求立方根的类似算法,有可能是从实际的立方体推导出来的,就像这里的算法是从正方形导出来的一样.中国人终于将这些思想发展成为一个能求解任意阶多项式方程的详尽的方法,这个方法我们将在第六章中来讨论.

1.8 毕达哥拉斯定理

巴比伦人求平方根的问题之一是与一正方形的边与对角线之间的关系相联系的.这个关系是毕达哥拉斯定理的一种特殊情况;毕达哥拉斯定理告诉我们:在任一直角三角形中,两直角边的平方和等于斜边的平方.这个定理,以公元前6世纪希腊哲学家和数学家的名字命名,可以有理由认为是数学中最重要的基本定理,因为它的推论和推广有着广泛的应用.虽然这样称呼,它可是古代文明中最古老的定理之一;实际上,有迹象表明,早在毕达哥拉斯之前至少一千年人们就已经知道这条定理了.

有些学者曾经论证,英国的与天文学有关的石庙在公元前第三个千年中建造时就用到了毕达哥拉斯定理,特别是毕达哥拉斯三数组的知识.毕达哥拉斯三数组是这样的一组三个数 (a,b,c),它们之间存在关系 $a^2+b^2=c^2$.不过这种说法根据并不充分.[22] 在巴比伦泥板普林顿(Plimpton)322号上有确切得多的证据表明了对毕达哥拉斯三数组的兴趣(图1.22),这块泥板的年代大约在公元前1700年.现存的这块泥板由四列数组成.(其它几列可能在左边断落部分).下面是这块泥板上的数,用现代十进制记号重新表达,最近的编辑者们作了少量的修正,并在其左端加了一列(第五列)推测数字.

y	$\left(\frac{x}{y}\right)^2$	x	d	#
120	0.983 4028	119	169	1
3456	0.949 1586	3367	4825	2
4800	0.918 8021	4601	6649	3

13 500	0.886 2479	12 709	18 541	4
72	0.815 0077	65	97	5
360	0.785 1929	319	481	6
2700	0.719 9837	2291	3541	7
960	0.684 5877	799	1249	8
600	0.642 6694	481	769	9
6480	0.586 1226	4961	8161	10
60	0.5625	45	75	11
2400	0.489 4168	1679	2929	12
240	0.450 0174	161	289	13
2700	0.430 2388	1771	3229	14
90	0.387 1605	56	106	15



图 1.22 普林顿 322(来源:乔治 · 阿瑟 · 普林顿藏品,哥伦比亚大学稀有图书和手稿图书馆).

对近代学者来说,首先要判定这不是一张陶器商业订货单,而是一件数学著作,然后还要给它一个合理的数学解释,这可是一项重大的数学侦查任务.但他们果真找到了一个解释.栏目为 x 和 d 的两列(这两个栏目原文可译为"宽方边" 和"斜方边"),每行均包含了毕达哥拉斯三数中的两个数.很容易从 x 列数的平方减去 d 列中同一行数的平方.每一次都会得出一个完全平方,它的平方根即重建的 y 列所示.最后,剩下的列表示商$\left(\frac{x}{y}\right)^2$.

这些三数组是怎样导出来的,又为什么要导出这些数?这么大的毕达哥拉斯三数组是无法用试

探法来找到的. $\left(\frac{x}{y}\right)^2$ 这一列的栏目意义不清,但有点像下面的意思:"矩形的辅助正方形,它被划出来从而宽度……",这对该表是怎样制作出来的问题多少给我们一点启示.为了求出方程 $x^2+y^2=d^2$ 的整数解,我们可以用 y 除全式,并先来求 $\left(\frac{x}{y}\right)^2+1=\left(\frac{d}{y}\right)^2$ 的解,或者,令 $u=\frac{x}{y}$,以及 $v=\frac{d}{y}$,那就是求 $u^2+1=v^2$ 的解.后一方程等价于 $(v+u)(v-u)=1$.就是说,我们可以把 $v+u$ 和 $v-u$ 看成面积为1的一个矩形的两边(图1.23),现在从这个矩形截出一条边长为 u 和 $v-u$ 的矩形,将它旋转90°并移至左下方,结果得到一个磬折形,它的两条长边都是 v,这个磬折形正好是两个正方形之差:$v^2-u^2=1$.注意那个"亏缺的"正方形,就是被"划掉"的那个,其面积为 $u^2=\left(\frac{x}{y}\right)^2$,恰是泥板上最左边的那一列中的项.为了计算泥板的项,首先取 $v+u$ 的值.下一步从倒数表中查得 $v-u$.解出 u 与 v,然后乘以一个适当的数 y,就得出整个毕达哥拉斯三数组.例如,若 $v+u=2;15(=2\ 1/4)$;其倒数 $v-u$ 为 $0;26,40(=4/9)$.解出 v 和 u 得 $v=1;20,50(=1\ 25/72)$ 和 $u=0;54,10(=65/72)$.将此二值乘以 $1;12=72$,则得 x 和 d 分别为65和97,这就是泥板中的第五行数字.反之,泥板第一行的 $v+u$ 可以这样来算:将 $169/120(=1;24,30)$ 与 $119/120(=0;59,30)$ 相加而得 $288/120(=2;24)$.

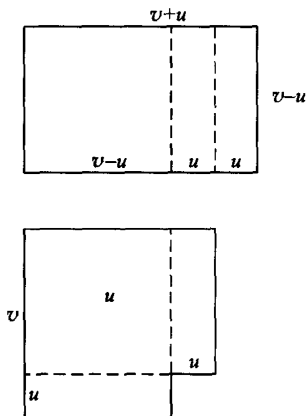


图 1.23　面积为 1 的矩形转变为两个正方形.

泥板上又为什么选择这些特殊的毕达哥拉斯三数组呢?我们同样也不能作出肯定的回答.但是如果我们计算泥板上每一行的 $v+u$,就会注意到,它们形成了一个由正则六十进数组成的递降数列,它们在 1;48 到 2;24 之间,不超过四位.不是所有这样的数都被包括进来了——有五个不见了——但这可能是书记员认为这个表没有它们就已经足够长了.他也可能讨论了大于2;24或小于1;48的数,只是记有这些内容的泥板至今尚未发掘出来.无论如何,他总算编出了一张整数的毕达哥拉斯三数组,用这些三数组可以设计训练学生的问题,而教师知道这些题会有整数解,或有有限的六十进制分数解.

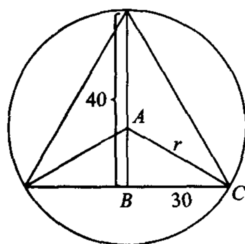


图 1.24　一个等腰三角形的外接圆.

不管这个方法是不是巴比伦的书记员用来编写普林顿泥板的方法,事实是这些书记员确实是知道毕达哥拉斯关系的.而且尽管这张特定的表除了各列的栏目名外,看不出有什么几何上的关系,但在一些老巴比伦泥板上有明确地将毕达哥拉斯定理应用到几何上去的问题.例如,泥板BM85196有这样一个问题,一根长为30的梁靠在墙上.上端向下滑了6的距离.下端移了多远?即 $d=30,y=24$ 已给,要求 x.书记员就用了毕氏定理来计算 x: $x=\sqrt{30^2-24^2}=\sqrt{324}=18$.在现代伊朗发现的一块泥板上有一个较为复杂一点的例子.问题是要求一等腰三角形的外接圆的半径,这个三角形的高为40,底为60.考虑直角三角形 ABC(图1.24),它的斜边为所要求的半径,书记员用毕达哥拉斯定理导出了方程 $r^2=30^2+(40-r)^2$.然后他算得 $1,20r=30^2+40^2=41,40$,再利用倒数算得 $r=(0;0,45)(41,40)=31;15$

在印度的《绳法经》中也提到了毕达哥拉斯定理和毕达哥拉斯三数组的例子,包括(5,12,13),(8,15,17),(7,24,25)和(12,35,37).还画了一张图来表示作一个正方形等于两个给定正方形的作图法,里面明确地采用了这个定理.在中国的文献资料中有关于这个定理的更详尽的内容,特别是

在《九章》一书的第九章中，这一章是专门讨论直角三角形问题的，在所有这些问题中都预先假定有毕达哥拉斯定理. 例如，在问题6中，已知一个正方形的池塘，边长为10尺，芦苇生在水中央，其顶高出水面1尺. 把芦苇拉到池岸，其顶刚刚碰到岸边，求水深和芦苇的长. 在图1.25中，$y = 5, x + a = d$，其中 $a = 1$. 一个现代的解可能是这样来开始的，令 $d^2 = x^2 + y^2$. 用简短的计算就可得出 $x = \frac{y^2 - a^2}{2a}$. 将所给的数值代入，得 $x = 12$，从而 $d = 13$. 《九章》中的解法是这样讲的："半池方自乘，以出水1尺自乘，减之，余，倍出水得水深. 加出水数，得葭长"[24]. 将这个算法译成公式就是我们已经推导出的 $x = \frac{y^2 - a^2}{2a}$. 然而有一点不清楚的是，到底《九章》的作者是用了上述代数解法，还是用了如图所示的几何解法，从图中可见有 $y^2 = AC^2 = AB^2 - BC^2 = BD^2 - EG^2 = DE^2 + 2CE \times BC = a^2 + 2ax$. 但有一点可以肯定的是，作者在使用毕达哥拉斯定理上是非常熟练的.

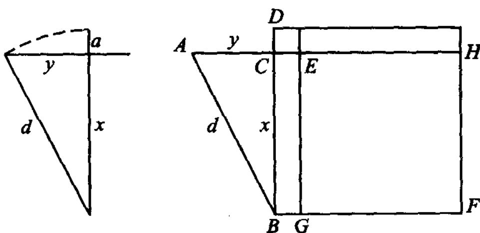


图1.25 池塘中芦苇的长度. 摘自《九章》第九章，问题6.

我们还要进一步指出，与《九章》第九章中的所有问题一样，问题6中的答数都是有理数. 因为在每个问题中都含有直角三角形，可由此推知，正如在巴比伦泥板教材中一样，问题就是特意这样来编的，使得这些直角三角形的各边长都为有理数. 不仅用到了(3,4,5)和(5,12,13)这样一些人们所熟知的三数组，还用到一些不太明显的三数组，如(55,48,73)和(91,60,109). 作者是怎样算出这些三数组而使他的问题能得出有理数答案的呢?

该章第14题给了我们一点线索. 现有 A, B 二人从同一地点出发. A 的速率为7，B 的速率为3. B 向东走；A 先向南走了10步，再向东北方向走，直至与 B 相遇. 问：A, B 各走了多远?作一个直角边为 x 和 y，斜边为 z 的直角三角形，我们有 $y = 10, z + y = \frac{7}{3}x$. 作者然后算得

$$z = \frac{7^3 + 3^3}{2}v, \quad y = \frac{7^3 - 3^3}{2}v, \quad x = 7 \cdot 3v.$$

式中 v 是一个任意常数. 由于 $y = 10$，v 必定为1/2，因而有 $z = 14\ 1/2$ 步，$x = 10\ 1/2$ 步作为所求的答案. 这里重要的一点是，作者表明了如何来一般地计算毕达哥拉斯的三组数. 因为如果 $z + y = \frac{a}{b}x$，以及 $z^2 - y^2 = x^2$，则有 $z - y = x^2/(z + y) = \frac{a}{b}x$，以及

$$z = \frac{a^2 + b^2}{2ab}x, \quad y = \frac{a^2 - b^2}{2ab}x.$$

由此得出公式

$$x = ab, \quad y = \frac{a^2 - b^2}{2}, \quad z = \frac{a^2 + b^2}{2}.$$

在 a, b 为奇数和 $a > b$ 的情况下，x, y, z 必定为一毕达哥拉斯的三数组. 例如，三数组(55,98,73)可通过取 $a = 11, b = 5$ 得到，而三数组(91,60,109)可以由取 $a = 13, b = 7$ 得到.

在求边为整数的直角三角形上，毕达哥拉斯三数组非常有用. 如果这确是这种三数组发展的理由，那么就可以认为毕达哥拉斯定理的几何意义是早已为人所知. 接下来很自然的问题就是，这个定理是怎么被发现和怎样"被证明"的. 和在许多其他领域中一样，没有对这一发现的记载. 在所有提到过的课本中都是把它当作已知的. 然而，关于"证明"倒是有迹可循.

印度的《绳法经》在它的构造圣坛的法则中给出了这样一个论断："一个正方形的两相对顶点

的连线所产生的面积是该正方形面积的二倍”.[25] 这个论断使人想到图 1.26(a),从这张图可以明显地看出该论断的证明来.由于这只是毕达哥拉斯定理的一个特例,可以猜想,普遍情形下的毕达哥拉斯定理可以通过改进这个图来证明.事实上,在至少是公元前数百年中国人的一本《周髀算经》中就确有这样一张改进的图(图 1.26(b)).附带有一段注文如下:

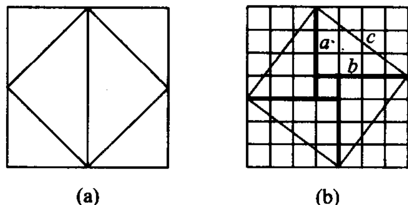


图 1.26 (a) 印度人对毕达哥拉斯定理的一个特例的证明;(b) 中国人对毕达哥拉斯定理的证明.

> 故折矩,以为句广三,股修四,径隅五.既方之,外半其一矩,环而共盘,得成三、四、五.两矩共长二十有五,是谓积矩.[26]

尽管注文和图形是对特殊三角形(3,4,5) 来讲的,(在最后两行中的) 证明却具有普遍意义.用 a 来记宽度,用 b 记长,用 c 记对角线长.则可论证如下:$(a+b)^2-2ab=c^2$;由于$(a+b)^2=a^2+b^2+2ab$,立即就得到了毕达哥拉斯定理.从几何上来看,论证直接有赖于用两种方式对大的正方形进行剖分,第一种是将它分成一个边长为 a 的正方形,一个边长为 b 的正方形,加两个面积为 ab 的矩形;而第二种剖分是将它分成一个边长为 c 的正方形加两个面积为ab 的矩形.同样立即可以推得所要的结论.

上述论证能否算是证明呢?要想满足现代的标准,还必须证明,内接的图形(边长为 c 的正方形) 和外接的图形(边长为 $a+b$ 的正方形) 的确是正方形.对古人来说,对今天大多数的学生来说可能也一样,这是显然的.中国人没有用一个公理体系来证明定理的观念.“证明” 只不过是意味着令人信服的论证.事实上,在希腊文中,定理(theorem) 这个字来自“theorein”,它表示“注视” 的意思.如果你注视着这个图形,你立即就会看出这个定理,巴比伦人对这个定理的最早记载也根本没有对它的论证.然而,这里所讲的论证,那些书记员们肯定也是能够做到.

1.9　二次方程

含有两个未知量的乘积或一个未知量的平方的问题会导出我们今天称为二次方程的方程.例如,涉及毕达哥拉斯定理的问题就通常会导致这样的方程.这种方程是古代巴比伦人的一个主要研究领域,它们也出现在中国的文献资料中.在这两个文明中,解题的方法是以几何概念为基础的,就是说,以几何的正方形和矩形这些概念,而不是以算术上的平方和乘积的概念为基础的.

中国的《九章》中的第九章,讲直角三角形的那一章,有几个问题,翻译过来就是二次方程.例如,问题 20 要求解方程 $x^2+34x=71\,000$.遗憾的是,该书的作者在这里只讲了解为 $x=250$,方法一点也没有讲.然而在第六章中有一个解这种问题的方法,它与先前讨论过的中国人求平方根的算法有密切的关系,这个算法有它的几何根源.换言之,它是一种递归算法,每一步得出一个更接近正确答案的近似.

《九章》中的多数二次方程问题翻译过来都是有两个方程的方程组问题.例如,问题 11 是这样一个问题:有一扇门,高度比宽度多 6.8 尺,两对角顶点间距为 10 尺,问门高和宽各是多少.应用毕达哥拉斯定理,这个问题就可译成下述方程组:

$$x-y=6.8,\qquad x^2+y^2=100.$$

中国人的解法是以先前已提出的对毕达哥拉斯定理的“证明” 为基础的.如果我们将这个问题重写

成更一般的形式:$x - y = d, x^2 + y^2 = c^2$,则在那个证明中所采用的图形表明有$(x+y)^2 = 4xy + (x-y)^2$以及$c^2 = 2xy + (x-y)^2$,因而有$4xy = 2c^2 - 2(x-y)^2$.由此推得$(x+y)^2 = 2c^2 - (x-y)^2$,即$x + y = \sqrt{2c^2 - (x-y)^2}$,或最终有

$$\frac{x+y}{2} = \sqrt{\frac{c^2}{2} - \left(\frac{d}{2}\right)^2}.$$

由此式确定的步骤使得作者能首先定出$x + y = 12.4$,然后,将此式与$x - y = 6.8$相结合,得出解为$x = 9.6, y = 2.8$.

尽管《九章》中还包括了其他一些可解释成线性和二次方程组的问题,但来自古代的大多数二次方程的例子都是出自巴比伦的文献.[27].事实上,许多古巴比伦泥板上列有大量的二次方程.某些这种巴比伦问题的标准形式

$$x + y = b, \quad x \cdot y = c$$

使人想到,这些巴比伦人原来是想处理一个矩形的面积与其周长之间的关系.看来在古代很多人以为一块田的面积只与它的周长有关.有各式各样的故事表明,那些懂得更多的人就利用这一点来占那些持有这种错误认识的人的便宜.于是不难想到,那些巴比伦的书记员为了说明周长相同的矩形可以有各种不同的面积,编制了一给定周长为$2b$,但长x与宽y各不相同的矩形面积c的表.研究这种表中各种长$x = \frac{b}{2} + z$和宽$y = \frac{b}{2} - z$与矩形面积$c = \left(\frac{b}{2} + z\right)\left(\frac{b}{2} - z\right) = \left(\frac{b}{2}\right)^2 - z^2$之间的关系,可以得出这样的结论,就是他们已经注意到了$z = \sqrt{\left(\frac{b}{2}\right)^2 - c}$,从而得知

$$x = \frac{b}{2} + \sqrt{\left(\frac{b}{2}\right)^2 - c}, \quad x = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}$$

是上述方程的解.总之,巴比伦的书记员们就是用我们这些现代公式所描述的算法来求解这类问题的.

自然,巴比伦人不会写出公式来.每个问题都是用在长度、宽度和面积上标记数字的方式提出,再指示特定的数值计算方法,这种方法我们可以用上述公式来解释.例如,拿泥板 YBC4663 上的一个问题$x + y = 6\ 1/2, x \cdot y = 7\ 1/2$来讲.书记员首先将 6 1/2 取半得 3 1/4.然后将 3 1/4 平方得 10 9/16.将它减去 7 1/2,余 3 1/16.取其平方根得 1 3/4.于是长度为 3 1/4 + 1 3/4 = 5,宽度则为 3 1/4 − 1 3/4 = 1 1/2.无论这个方法的最终根源何在,仔细阅读泥板上的文字似乎表明,书记员的脑子里已经有一个几何方法(图 1.27),为了具有普遍性,其中各条边已经按照一般方程$x + y = b, x \cdot y = c$来作标记.[28] 书记员开始将和b取半,然后再作这一半的平方.由于$\frac{b}{2} = x - \frac{x-y}{2} = y + \frac{x-y}{2}$,$\frac{b}{2}$的平方比原矩形的面积$c$要大过一个$\left(\frac{x-y}{2}\right)$的平方,即

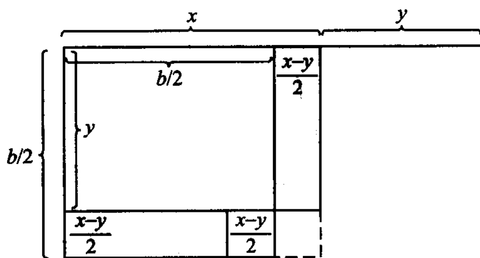


图 1.27 解方程$x + y = b, x \cdot y = c$的几何方法.

$$\left(\frac{x+y}{2}\right)^2 = xy + \left(\frac{x-y}{2}\right)^2.$$

那么由图 1.27 可见,如果将这个正方形的边,即$\sqrt{\left(\frac{b}{2}\right)^2 - c}$,加到$\frac{b}{2}$上,就得到了长度$x$,如果从$\frac{b}{2}$

减去这个值,就得宽度 y,因此这个算法正好就是上一节中所指出的方法.

对巴比伦人为其他类型的二次方程问题所研究出来的算法可以作类似的几何解释.例如,方程组

$$x - y = b, \qquad x^2 + y^2 = c$$

可以用下述现代公式所表述的算法来求解.

$$x = \sqrt{\frac{c}{2} - \left(\frac{b}{2}\right)^2} + \frac{b}{2}, \quad y = \sqrt{\frac{c}{2} - \left(\frac{b}{2}\right)^2} - \frac{b}{2},$$

尽管这个问题和《九章》中一个关于门的问题形式相同,看来巴比伦人是用了一种不同的几何概念来求它的解.图 1.28 表明有

$$x^2 + y^2 = 2\left(\frac{x+y}{2}\right)^2 + 2\left(\frac{x-y}{2}\right)^2.$$

由此得出 $c = 2\left(\frac{x+y}{2}\right)^2 + 2\left(\frac{b}{2}\right)^2$,从而有$\frac{x+y}{2} = \sqrt{\frac{c}{2} - \left(\frac{b}{2}\right)^2}$.因为 $x = \frac{x+y}{2} + \frac{x-y}{2}$,以及 $y = \frac{x+y}{2} - \frac{x-y}{2}$,即得要求的结果.

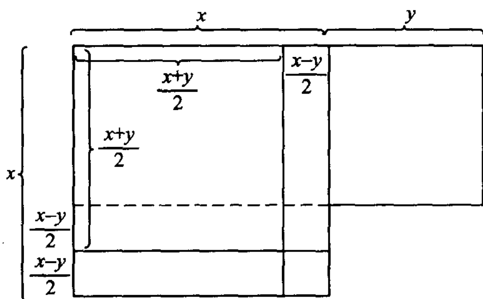


图 1.28　求解方程组 $x - y = b, x^2 + y^2 = c$ 的几何方法.

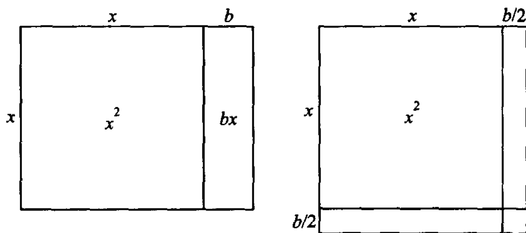


图 1.29　方程 $x^2 + bx = c$ 根的公式的几何意义.

除了方程组,巴比伦人也求解单个的方程.在泥板 BM13901 上就有这样的问题,其中有一个如下:一正方形的面积与它的一条边的 4/3 之和为 11/12.求其边长.用现代的术语来说,要解方程 $x^2 + (4/3)x = 11/12$.为了求解,书记员告诉我们取 4/3 的一半,得 2/3,取 2/3 的平方,得 4/9,然后将这个结果加到 11/12 上去,得 1 13/36.这个值是 7/6 的平方.从 7/6 中减去 2/3 得 1/2,这就所需要的边长.这个巴比伦的算法可以容易地翻译成现代解方程 $x^2 + bx = c$ 的公式,即

$$x = \sqrt{\left(\frac{b}{2}\right)^2 + c} - \frac{b}{2}.$$

可以认得出来这就是二次方程求根公式的一种形式.然而,问题是巴比伦人是怎样来阐释他们的方法的呢?初看起来,这个问题的陈述不具有几何性质,因为要我们做的是将一边的一个倍数加到一个面积上去.但解的几何解释似乎提示我们,这个倍数可以看成是一个长度为 x,宽为 4/3 的矩形,这个矩形要加到边长为 x 的正方形上去(图 1.29).在这个解释之下,解法就相当于将这个拼在该正方形某一边上的矩形切成两半,并将切下的一半挪到这个正方形的底部.加上边长为 $b/2$ 的正方形就可以将所述图形"补成正方形".于是未知长度 x 显然等于这个新的正方形的边长与 $b/2$ 之差,完全与公式所讲的一样.

对于巴比伦人得出方程 $x^2 - bx = c$ 的解 $x = \sqrt{(b/2)^2 + c} + \frac{b}{2}$ 的算法，其几何意义也可以类似地来说明.不过，应该记住，根的"二次式"对巴比伦的书记员们来说，它的意义和我们今天所指的不一样.首先，书记员们在解这两种类型的方程 $x^2 + bx = c$ 或 $x^2 - bx = c$ 时所用的算法是不同的，因为这两类问题不一样；它们有不同的几何意义.另一方面，对一个现代的数学家来说，这些问题却是一样的，因为 x 前的系数可以取正数，也可以取负数.其次，在这两种情况下现代的二次式对每个方程都会给出一个正的解和一个负的解，然而，负的解没有几何意义，因而被巴比伦人所完全忽视.有趣的是，有两个正根的，形式为 $x^2 + c = bx$ 的方程也被巴比伦人忽略了.在泥板中没有出现过这种方程.尽管能把方程写成这种形式的问题的确存在，书记员们显然想不到，单个方程居然能使同一未知量获得两个不同的值.所以他们想尽一切办法阻止这种可能性出现.做到这一点的最简单的办法就是把问题重新写成这样的形式：$x + y = b, x \cdot y = c$，这样两个解分属两个未知量.方程 $x^2 + c = bx$ 只在 $\left(\frac{b}{2}\right)^2 = c$ 这个惟一特殊的情况下求解过，因为在这种情形下只有一个解.

有很多巴比伦的泥板上有许多分成组的二次方程的问题.这种一组一组的问题代数上常常很复杂.有些问题组是以"抽象"的形式写出来的，另一些则从前后文来看似乎是"真实世界"中的问题，例如工程中的问题.这后一类二次方程的问题实际上是人为编制的，和我们在当今代数教科书中所遇到的有些问题是一样的.这些作者也知道这些问题是编造的，这一点从给定一组中所有问题有相同的解这一典型情况就可以看出来.所以看来这些泥板是用来培养解题技巧的.换言之，解各种问题的目的不是为了确定答案是什么，而是为了学习把复杂问题约化为简单问题的各种方法.因此，可以这样设想，一般的数学泥板，特别是这些二次方程泥板，是用来培训那些未来国家的领袖人物的大脑的.换言之，解二次方程并不是在实际上有那么重要 —— 没有多少实际情况需要它们.重要的是用来培养学生解决一般问题的能力，一种用来处理一个国家领导人需要解决的日常问题的能力.这种能力不仅包括能够实施已经开发出来的算法，而且包括懂得如何和在何时去修正这些方法，以及如何把比较复杂的问题归结为一些已经解决了的问题.今天的学生常常被告知，学习数学是为了"训练大脑".看来在过去的四千年来老师们一直在教导他们的学生这同一条真理.

习　题

计数

1. 确定在你和你的同学所知的语言中数 18 和 40 是用什么词来表示的.比较这些词汇的构造.有没有任何本质上不同于补遗 1.1 中的那些词汇的形式？
2. 将 125 用埃及象形文字和巴比伦楔形文字表示出来.
3. 在至少大约公元前 450 年，希腊人用了一个以他们的字母为基础的数码系统.表示法如下：

α	1	ι	10	ρ	100
β	2	κ	20	σ	200
γ	3	λ	30	τ	300
δ	4	μ	40	υ	400

ϵ	5	ν	50	ϕ	500
ϛ	6	ξ	60	χ	600
ζ	7	o	70	ψ	700
η	8	π	80	ω	800
θ	9	ϙ	90	ϡ	900

其中代表 6 的字母ϛ(digamma),代表 90 的字母 ϙ(koppa),和代表 900 的字母ϡ(Sampi)现在已不再使用了.因而 754 是写成 $\psi\nu\delta$,而 293 则写成 σϙγ.为了表示千,就在字母 α 到 θ 的前面作个记号;例如,θ 表示 9000.再更大的数就是 M 表示万(myriads,10 000),万的个数就写在 M 的右上角:M^{δ} = 40 000,$M'^{\zeta\rho o\epsilon}$ = 71 750 000.试用希腊字母记号表示 125,62,4821 和 23 855.

4.中国商代所用的表示数的基本符号是

1	2	3	4	5	6	7	8	9	10	100	1000
一	二	三	亖	☒	介	十	)(	𠄎	丨	⊚	𠂉

对 20,30,40 有组合符号(即,U Ш Ш),但一般来说,记号是按本书正文中所讲过的方案来组合的.例如,88 是)()(,162 是⊚介二.写出下述各数的中国记法:56,554,63 和 3282.

算术计算

5.用埃及人的算法做 84 被 5 除的除法

6.用埃及人的乘法做 $7\overline{2}\,\overline{4}\,\overline{8}$ 乘以 $12\,\overline{\overline{3}}$.注意,被乘数的每一项都要分别乘以 $\overline{\overline{3}}$.

7.《兰德数学纸草书》中除以 2 的除法表的一部分如下:$2 \div 11 = \overline{6}\,\overline{66}$,$2 \div 13 = \overline{8}\,\overline{52}\,\overline{104}$,$2 \div 23 = \overline{12}\,\overline{276}$. $2 \div 13$ 的计算如下:

1	13
$\overline{2}$	$6\,\overline{2}$
$\overline{4}$	$3\,\overline{4}$
$\overline{8}$	$1\,\overline{2}\,\overline{8}$ '
$\overline{52}$	$\overline{4}$ '
$\overline{104}$	$\overline{8}$ '
$\overline{8}\,\overline{52}\,\overline{104}$	$1\,\overline{2}\,\overline{4}\,\overline{8}\,\overline{8}$
	2

对 2 除以 11 和 23 作类似的计算以检验其结果.

8.如上所述给定 $2 \div 13 = \overline{8}\,\overline{52}\,\overline{104}$,则 3,4,5,…,12 被 13 来除的单位分数就很容易算得.例如 $3 \div 13 = \overline{8}\,\overline{13}\,\overline{52}\,\overline{104}$(因为 $3 = 1 + 2$),又 $4 \div 13 = \overline{4}\,\overline{26}\,\overline{52}$(因为 $4 = 2 \times 2$).类似地来计算 $5 \div 13$,$6 \div 13$ 和 $8 \div 13$.

9.《兰德纸草书》的问题 79 中的第二部分是这样的:

房子	7
猫	49
老鼠	343
谷穗	2 401
赫卡特	16 807
总计	19 607

有人猜想这个问题有点像一首古老英格兰童谣"在我去圣艾弗斯的路上".于是整个问题可以这样来解读:"一座庄园里有 7 幢房子,每幢房子里有 7 只猫.每只猫逮住 7 只鼠,每只鼠要吃 7 棵穗,每棵穗能产生 7 赫卡特 * 的粮食.在这个庄园里总共有多少东西?".这个问题的第一部分证明了,2801 乘以 7 的积为 19607,证明这个答案是正确的,因为几何级数 7 + 49 + 343 + 2401 + 16 807 的和等于这个值.

10. 通过基数为 60 的除法证明,1 ÷ 7 将得出六十进制的循环小数 0;8,34,17,8,34,17….

11. 求以 60 为基数的 18,32,54 和 64(= 1,04) 的倒数. n 应满足什么条件才能保证它的倒数为有理的六十进制小数——即有限的六十进制小数?

12. 在巴比伦数系中,作 25 乘以 1,04 和 18 乘以 1,21 和乘法.再用倒数表作 50 除以 18 和 1,21 除以 32 的除法.采用我们的、经修正后适用于以 60 为基数的标准乘法.

13. 一个数与它的 1/7 相加后为 19,这个数是什么数?用假位法求解.

14.《兰德纸草书》的问题 72 是:"用 100 个 10 比苏(pesu)** 的面包去交换 45 比苏的面包,能换多少?" 这个问题的解法如下:"求出 45 超过 10 的数量,为 35.将此 35 除以 10,得 $3\bar{2}$.将 $3\bar{2}$ 乘 100,结果为 350.将此 350 加上 100,得 450.于是可以说 100 个 10 比苏的面包可以换 450 个 45 比苏的面包".[29],将此解法用现代术语来表达.将此解法与书中问题 75 的解法相比较.这个解是怎样显示出"线性"性质的?

15. 解《九章》第三章问题 3:三人各有钱 560,350 和 180.共付税款 100,按各人所有钱数按比例分配税款.各应付多少?

16. 用中国人的方法解《九章》第八章第三题:现有上等禾 2 捆,中等禾 3 捆,下等禾 4 捆,各自的实都不满一斗.如果三种禾分别依次借取中等禾,下等禾,上等禾各一捆,那么实都刚满一斗,问:三种禾每捆的实各有多少?说明用中国的解法要涉及到负数的应用.

17. 解《九章》第七章第一题:数人合购一物,每人出 8,盈 3;每人出 7,不足 4;问:人数、物价各是多少?用盈不足法求解.

18. 解《九章》第六章第二十六题:现有水池,用五条渠道灌水.已知单独灌时,开第一条渠,1/2 日把池灌满;开第二条渠,1 日把池灌满;开第三条渠,1/3 日把池灌满;开第四条渠,3 日把池灌满;开第五条渠,5 日把池灌满.如五条渠同时开放,多少日把池灌满?(这个问题是我们所知这类问题中最早的一个.类似的问题也出现在较晚的希腊,印度和西方的教材中.)

19. 解《九章》第六章第二十八题***:现有人携大米上路,经过了三个收税站.在第一站他上缴了所有大米的 1/3,在第二站,他上缴了余下的 1/5.而在第三站又上缴了这余下的 1/7.在通过所有这三站后,他只剩下 5 斗米了.问:他原来带多少大米出发?(这个问题的各种不同版本出现在后来各种文明的文献中.)

初等几何

20. 已给一半径为 1 的圆和一条中心角为 90° 的弦.证明该弦长 s 为$\sqrt{2}$,"矢"长为$\frac{2-\sqrt{2}}{2}$.计算由此弦所确定的弓形的面积,分别用书中的中国公式和现代的方法来计算,并比较其结果.再对中心角为 60° 和 45° 的弓形作相同的计算.

21. 证明巴比伦人的"船形"(见图 1.14) 面积公式为 $A = (2/9)a^2$,式中 a 为船形的弧长(四分之一圆周).再证这个船形长横截线长度为$(17/18)a$,而短横截线的长度为$(7/18)a$.利用巴比伦人的圆面积公式 $C^2/12$ 和$\sqrt{2}$ 的数值 17/12.

22. 证明巴比伦人的"牛眼"(见图 1.14) 面积公式为 $A = (9/32)a^2$,式中 a 为其弧长(三分之一圆周).再证它的长横截线的长度为$(7/8)a$,短横截线长为$(1/2)a$(采用巴比伦人的圆面积公式 $C^2/12$ 和$\sqrt{3}$ 的值 7/4).

23. 对埃及人的圆面积公式 $A = \left(\frac{8}{9}d\right)^2$($d$ 为圆的直径) 的推导已有各种推测. 其中之一采用圆片计数法(circular

---

* 见本书 §1.4—— 译者注.
** 见本书 §1.4—— 译者注.
*** 应为第 27 题 —— 译者注.

counters),据说古埃及人用过这种方法.用硬币做实验,譬如说,取硬币的直径为1,证明一个直径为9的圆基本上可用64个直径为1的圆填满.(首先将一个硬币放在中心;再用六个硬币围绕着它排成一圈.依此类推.)64个直径为1的圆也能填满边长为8的正方形,利用这一明显的事实来说明,埃及人可能是怎样导出他们的公式的.[30]

24. 就《莫斯科纸草书》中的截棱锥体,将书中给出它的正确体积与用巴比伦的错误公式 $V=\frac{1}{2}(a^2+b^2)h$ 所算出的体积相比较.算出误差的百分率.对一底面积为10,顶面积为8,高为2的截棱锥体作同样的计算.

25. 在印度的《绳法经》中,祭司们给出了一个求一个面积等于一给定正方形的圆的算法.在正方形 $ABCD$ 中,令 M 为它的两条对角线的交点(图1.30).以 M 为圆心,以 MA 为半径作一个圆;令 ME 为这个圆的一条与边 AB 垂直的半径,它交 AD 于点 G.令 $GN=\frac{1}{3}GE$.则 MN 为所求圆的半径.证明:如 $AB=S, 2MN=d$,则 $\frac{d}{s}=\frac{2+\sqrt{2}}{3}$.,并说明这意味着 π 等于3.0883.

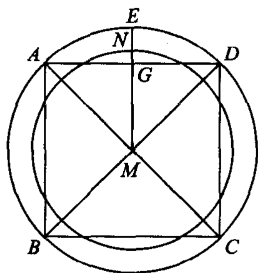


图1.30　"化圆为方"的印度算法.

天文计算

26. 在一本天文学课本上查找月球升起位置的18.6年的周期.讨论它的天文学基础和我们怎样用它来预测月食.你认为斯通亨格的祭司们是不是有可能作过这种预测?

27. 在中国古代历书上查找他们是怎样调和月球与太阳的运动周期的细节.

28. 查找一本有犹太历法细节的参考书.写一篇简短的报告,谈谈每年的长度是怎样确定下来的.

平方根

29. 将 $\sqrt{2}$ 的巴比伦近似值1;24,51,10转化为十进小数,并确定这一近似的准确度.

30. 用假设是巴比伦人的求平方根的算法,从2这个值开始计算.证明有 $\sqrt{3}=1;45$.求出1;45的倒数到六十进制的第三位,并利用它来计算 $\sqrt{3}$ 到六十进制的第三位.

31. 证明 $12\ \bar{\bar{3}}\ \overline{15}\ \overline{24}\ \overline{32}$ 是 $\sqrt{164}$ 的一个好的近似.(这个值出现在一本新近发现的希腊 — 埃及纸草书上.)

32. 利用中国人的求平方根的算法(这与我们现今所教的方法是一样的)导出 $\sqrt{2}$ 的一近似十进序列,准确到五位数,再用公式 $\sqrt{a^2\pm b}=a\pm\frac{1}{2}b\left(\frac{1}{a}\right)$ 作相同的工作.比较这两种做法哪一种好.(即哪种算法能用最少量的计算得出精确到五位小数的近似?)

毕达哥拉斯定理

33. 试证如果取 $v+u=1;48(1\ 4/5)$,就可得普林顿322泥板上的第15行;如取 $v+u=2;05(2\ 1/2)$,就会得第9行.求解得出该泥板上第6行和第13行时的 $v+u$ 的值.

34. 普林顿322泥板的书记员在制作这个泥板时没有采用 $v+u=2;18,14,24$ 和它的倒数 $v-u=0;26,02,30$.试求与这两个值相联系的最小的毕达哥拉斯三数组.

35. 解《九章》第九章的第八题:墙高10尺,木杆斜靠着墙,杆顶齐墙顶.如杆脚向外移1尺,杆顶恰好落地,问:木杆长是多少?

36. 证明在《绳法经》一书中所给出的,作一个正方形等于两个正方形之差的作图法是正确的(图1.31):令 $ABCD$ 为大正方形,其边长为 a,而 $PQRS$ 为较小的正方形,边长为 b.从 AB 切出一段 $AK=b$,并作 KL 垂直于 AK,交 DC 于 L.以 K 为圆心,以 KL 为半径画圆弧交 AD 于 M,则以 AM 为边的正方形即为所求.

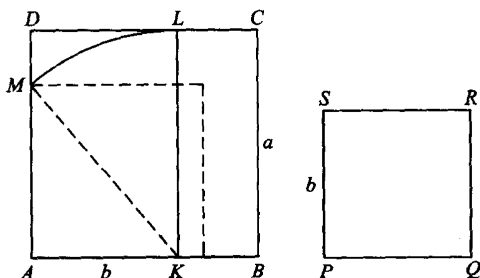


图 1.31　《绳法经》决定一个正方形等于两正方形之差的作图法.

二次方程

37. 求解老巴比伦泥板 BM13901 中的一个问题:第二个正方形的一边是第一个的 2/3 加 5,求它们的边长.
38. 求解《柏林纸草书》中的一个问题:如果一个面积为 100 平方肘尺的正方形等于面积较小的两个正方形的和,而且如果其中一个的边长为另一个边长的 $\bar{2}\,\bar{4}$(= 3/4) 倍,求这两个未知正方形的边长.
39. 解《九章》第九章第二十题:"现有正方形城,不知大小.每边正中开门,出北门 20 步有树,出南门走 14 步远后,折向西走 1775 步,见树,问:城的边长是多少?
40. 对巴比伦人解方程 $x^2 - bx = c$ 的"二次公式" 给一个几何的论证.
41. 考虑从一本古巴比伦教本上选择的方程组:

$$x = 30, \qquad xy - (x - y)^2 = 500.$$

证明将第一个方程代入第二个方程后会得出一个 y 的二次方程,这个方程有两个正根,这是巴比伦人没有研究过的类型.证明,当第一个方程的平方减去第二个方程时就得$(x - y)$的一个二次方程:$(x - y)^2 + 30(x - y) = 400$,它只有一个正根.
42. 解下述巴比伦人的问题:

$$x + y = 5\frac{5}{6}, \quad \frac{x}{7} + \frac{y}{7} + \frac{xy}{7} = 2$$

的办法是,先将第二个方程乘以 7,再从它减去第一个方程,从而将这个方程化成了标准形式.

讨论题*

43. 从表达式的简洁性,所需的记忆量,和算术计算的难易度等方面,讨论编组数系,符号数系和位值制数系(均以 10 为基数) 的优劣.
44. 仍以上述三种比较优劣的标准比较以 10 为基数和以 60 为基数的系统.是否有一个比上述两种基数都要好的基数?试说明之.
45. 设计一堂教位值制的课,采用巴比伦数系,以及特别是图 1.8 的乘法表.
46. 设计一堂用类似于(据信是) 巴比伦人采用过的几何论证方法来讲二次公式的课.
47. 从使用的方便程度上比较标准的除法和巴比伦人用倒数的方法.
48. 设计一堂根据埃及和中国史料中的例子来说明线性化思想的课.
49. 设计一堂课来向学生证明圆面积公式 $A = \pi r^2$ 的正确性,这里 π 定义为$\dfrac{C}{d}$,π 的值怎样来确定?你怎样来使学生信服,对任一圆,它的周长与半径之间的比率是一个常数?
50. 为什么要有这样一种历法,它的月份由月球运动的周期来确定?这种历法用途何在?今天我们没有这种历法会

* 要求"设计一堂课" 的讨论题的意图是帮助未来(各级) 教师在数学教学中应用历史.在本书的大多数习题组中都有这种问题.它对未来的教师在课堂上讲这种课程时可能会有用.

有什么损失?

51. 用中国的史料设计一堂讲毕达哥拉斯定理的课.

文献和注解

关于我们所讨论的古代文明的基本信息可以从,例如,下述文献中获得:William McNeill, *The Rise of the West*(Chicago, University of Chicago Press, 1970) 和《*Peoples and Places of the Past*》(Washington:国家地理学会,1983).叙述巴比伦数学的标准著作有:Otto Neugebauer,《*The Exact Sciences in Antiquity*》(Princeton;普林斯顿大学出版社;1951;New York:Dover,1969) 和 B. L. van der Waerden,《*Science Awakening*》(New York:牛津大学出版社).最近的研究论文有 Jens Hφyrup:《数学,代数和几何》,载 David N. Freedman 所编之 *Anchor Bible Dictionary*(New York: Doubleday, 1992, vol. Ⅳ. pp. 601 - 612) 以及 Jöran Friberg,"*Mathematics*",载 *Reallexikon der Assyriologie* 7(1987—1990)531 - 585(英文).巴比伦泥板的翻译和分析主要见 Otto Neugebauer: *Mathematische Keilschrift - Texte*(New York: Springer, 1973, 1935 原版的影印本). Otto Neugebauer 与 Abraham Sachs 的《*Mathematical Cuneiform Texts*》(New Haven:美国东方学会,1945) 以及 Evert Bruins and M. Rutten, *Textes mathématiques de Suse*(Paris: Paul Geuthner, 1961).有关埃及数学的最好来源是 Richard J. Gillings, *Mathematics in the Time of the Pharaohs*(Cambridge: MIT Press, 1972).还可参阅 Gillings,《*The Mathematics of Ancient Egypt*》, *Dictionary of Scientific Biography*(New York, Scribners, 1978), vol 15, 681 - 705. Arnold B. Chace 编辑的《兰德数学纸草书》是该文献目前可用的版本(Reston, Va:(N.C.T.M 国家数学教师委员会), 1967).(本书为美国数学学会 1927 和 1929 出版的原书的简编).对中国数学的一般性讨论见 J. Needham(李约瑟), *Science and Civilization in China*(Cambrige:剑桥大学出版社,1959)vol. 3 及李俨与杜石然的《中国数学简史》,英译本由 John N. Crossley 与 Anthony W. C. Lun 合译(Oxford Clarendon Press, 1987).对中国数学的一个简短的研究有 Frank Swetz 的"数学在古代中国的发展",载《数学杂志》52(1979)10 - 19. Frank Swetz 与 T. I. Kao 翻译了《九章算术》的第九章,并以 *Was Pythagoras Chinese*?的书名发表(Reston Va: N.C.T.M, 1977). Lay Yong Lam(兰丽蓉)的《九章算术:概览》,载 *Aichive for History of Exact Sciences* 47(1994)1 - 51,以及一本完整的德文译本:*Neun Bücher arithmetischer Technik*, Kurt Vogel 译(Braunschweig: F. Vieweg und Sohn, 1968).对古代印度数学的历史没有全面的论著,但有两本接近这方面的著作,B. Datta 和 A. N. Singh, *History of Hindu Mathematics*(Bombay: Asia Publishing House, 1961)(1935—38 原书的影印本)和 C. N. Srinivasiengar, *The History of Ancient Indian Mathematics*(Calcutta: The World Press Private Ltd, 1967).有一本更全面讲述印度科学的著作,包含了许多有关印度数学的有趣材料,是 Debiprasad Chattopadhyaya 的《*History of Science and Technology in Ancient India-The Beginnings*》(Calcutta: Firma KLM Pvt. Ltd, 1986).最后,比较地讨论这些古代社会的数学同时还讨论了这些数学可能的单一起源和传播问题的一本著作是 B. L. van der Waerden 所著的 *Geometry and Algebra in Ancient Civilizations*(New York: Springer, 1983).

1. Chace, *Rhind Mathematical Papyrus*, p. 27.
2. 关于数字的更多的信息,标准的参考文献为 Karl Menninger, *Number Words and Number Symbols*(Cambridge, MIT Press 1977).
3. Alexander Marshack 的 *Roots of Civilization* 第二版(Mount Kisco, NY: Moyer Bell Limited, 1991)一书对此有讨论.
4. 关于埃及建筑中零的使用见 Dieter Arnold, *Building in Egypt*, (New York: Oxford University Press, 1991), p. 17 和 George Reisner, Mycerinus: *The Temples of the Third Pyramid at Giza*(Cambridge: Harvard University Press, 1991). p, 76 - 77. 关于埃及会计中零的使用,见 Alexander Schalff, "Ein Rechnungsbuch des Königlichen Hofes aus der 13. Dynaste" 载 *Zeitschrift für Ägyptische Sprache and Altertumskunde* 57(1922), 58 - 59.
5. 见 Denise Schmandt-Besserat, *Before Writing: from Counting to Cuneiform*(Austin: University of Texas Press, 1992),该书中对此有广泛的讨论,并有很多插图,展示了在上一世纪(指 19 世纪)发掘出来的大量巴比伦证券文书.
6. 见 Théophile Obenga, L'Afrique dans L'Antiquité: Egypte Pharaonique-Afrique Noire (Paris: Présence Africaine, 1973),第 9 章.
7. 见 Gillings, *Mathematics in the Time of the Pharaohs*, pp. 45 - 50,另见 Wilbur Knorr, "*Techniques of Fractions in Ancient*

Eqypt and Greece",载 *Historia Mathematica*(1982),133 - 171.

8. Chace, *Rhind Mathematical Papyrus*, p.69.

9. Gillings,, *Mathematics in the Time of the Pharaohs*, p.173.

10. Neuegebauer and Sachs, *Mathematical Cuneiform Texts*, p.101. 泥板前的记号表示其保存处. 例如 YBC 表示 Yale Babylonian Collectlon, BM表示 British Museum, VAT表示 Vorderasiatische Abteilung, Tontafeln of the Berlin Museum, 而 Plimpton 则指 Plimpton Collection of Columbia University.

11. Vogel, *Neun Bücher*, p.78.

12. 三上义夫(Yoshio Mikami), *Mathematics in China and Japan*(New York:Chelsea,1974)(1913原书的重印),p,18.本书也提供了一个对中国数学的很好的,虽说是有点过时的研究.对研究日本数学来说,这还是惟一能得到的英文资料.

13. Vogel, *Neun Bücher*, p.82.

14. Gillings, *Mathematics in the Time of the Pharaohs*, p.139.进一步的分析,见 Hermann Engels,"*Quadrature of the Cizcle in Ancient Egypt*",载 *Historia Mathematica* 4(1977),137 - 140.

15. Abraham Seidenberg,"*The Rieual Origin of Geometry*",载 *Archive for History of Exact Sciences* Ⅰ(1962)488 - 527,p.515. 该文对古代印度的几何与其他地方的几何的关系有许多有趣的叙述. 还有,见 Seidenberg,"*On the Aiea of a Semicizcle*", *Archive for History of Exact Sciences* 9(1973),171 - 211 以及"*The Origin of Mathematics*", *Archive for History of Exact Sciences* 18(1978),301 - 342.

16. Gillings, *Mathematics in the Time of the Pharaohs*, p.188.

17. 这里对谷堆及相关图形体积的分析是根据 Eleanor Robson 的 Mesopotamian Mathematics:2100—1600 *B.C.*: Technical Constants in Bureaucracy and Education(Oxford: *Oxford University*, 1998). Robson 博士是今天巴比伦数学的一位比较活跃的研究者.在未来几年她将在这一领域发表许多著作,全都值得一读.

18. 对于斯通亨格的天文学,参见 Euan W. Mackie, *Science and Society in Prehistoric Britain*(London, Paul Elek, 1977).也可以参见 Gerold Hawkions, *Stonehenge Decoded*(New York, Dowbleday, 1965) 和 Fred Hoyle, *On Stonehenge*(San Francisco; Freeman, 1977).

19. 对此有一个有趣的虚构例子,见马克·吐温的《亚瑟王法庭上的一个来自康奈狄克州的美国佬》.上一注释中提到的 Hawkins 和 Hoyle 的书对这个 18.6 年的周期及其在月食预测上的应用有广泛的讨论.

20. Needham, *Science and Civilization*, p.406.

21. 见 Wang Ling and J. Needham,"*Hornei's Method in Chinese Mathematics: Its Origins in the Root Extraction poceduces of Han Dynasty*" 载《*T'oung Pao*》43(1955),345 - 401.

22. 关于这些圆石阵中毕达哥拉斯定理的讨论,见 Van der Waerden, *Geometry and Algebra*. 还参阅 Frank J. Swetz 在 *Historia Mathematica* 13(1986),83 - 85 中所作的评论以及, W.R. Knorr,"*The Geometec and the Archaeoastconomec*", *British Journal of the History of Science* 18(1985),202 - 211,后一篇论文强烈反对这样的观点,即圆石阵的建造者知道毕达哥拉斯定理.

23. 这里所讲对 Plimpton322 的分析是摘自 Eleanor Robson 的一个建议(见注释 17). 参考 Jöran Friberg,"*Methods and Traditions of Babylonian Mathematics I: Plimpton 322, Pythagorean Triples, and the Babylonian Triangle Parameter Equations*", *Historia Mathematica* 8(1981),277 - 318 和"Mathematik", *Reallexikon der Assyrologie*7(1987—1990),531 - 585,这里可以读到有关这块泥板上编制这个表的其他可能方法的详细讨论以及有关出现误差的可能原因. 又参阅 R.C. Buck,"Sherlock Holmes in Babylon" *American Mathematical Monthly*, 87(1980),335 - 345.

24. Swetz, *Was Pythagoras Chinese?*, p.30.

25. Jerold Mathews,"A Neolithic Oral Tradition for the Van der Waerden/Seidenberg Origin of Mathematics", *Archive for History of Exact Sciences* 34(1985),193 - 220, p.203.本文讨论了中国人和巴比伦人应用毕达哥拉斯定理和其他几何概念的情况.

26. Needham, *Science and Civilization*, p.22.

27. 关于巴比伦泥板中二次方程的详细讨论可参阅:Solomon Gandz,"The Origin and Development of the Quadratic

Equations in Babylonian, Greek, and Early Arabic Algebra" *Osiris* 3(1937)405 – 557 和 Solomon Gandz, "Studies in Babylonian Mathematics Ⅲ: Isoperimetric Problems and the Origin of the Quadratic Equation" *Isis* 32(1947), 103 – 115. 又参阅 Philip Jones "Recent Discoveries in Babylonian Mathematics", *Mathematics Teacher*, 50(1957), 162—165, 442—444, 570 – 571.

28. 关于巴比伦方法的几何基础有一个新鲜的见解，可参阅 Jens Hфyrup. "*Algebra and Naive Geometry: An Investigation of Some Basic Aspects of Old Babylonian Mathematical Thought*", *Altorientalische Forschungen* 17(1990). 本节中提到的若干几何概念均采自该著作.

29. Gillings, *Mathematics in the Time of the Pharaohs*, p. 134.

30. 这个想法来自 Paulus Gerdes, Three Alternate Methods of Obtaining the Ancient Egyptian Formula for the Area of a Circle, *Historia Mathematica* 12(1985)261 – 267. 该文还讲述了另外两种可能的方法.

古代社会数学成就概览

埃及 —— 约公元前 1800 年

单位分数　　线性方程
圆的测量　　阳历和阴历
截棱锥体　　锥体的体积(可能)

巴比伦 —— 约公元前 1700 年

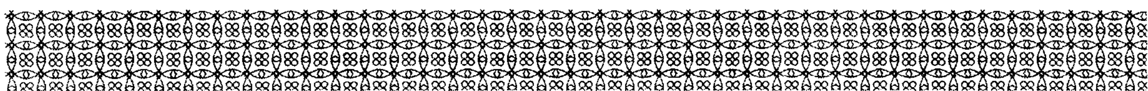
60 进位制系统　　两个线性方程的方程组
圆的测量　　二次方程和方程组
阳历与阴历　　平方根和立方根
平方根计算　　毕达哥拉斯定理
毕达哥拉斯三数组　　锥体的体积(可能)

印度 —— 约公元前 500 年

毕达哥拉斯定理　　圆的测量
平方根计算

中国 —— 约公元前 200 年

以 10 为基数的算筹　　少于五个方程的方程组
圆的测量　　锥体的体积
平方根与立方根的计算　　毕达哥拉斯定理
毕达哥拉斯三数组　　二次方程和方程组



第2章 希腊数学的开始

> 泰勒斯首先从希腊来到埃及,然后将几何研究引进希腊,他本人发现了许多命题,并指导他的学生研究那些可以推出其它命题的基本原理,他的一些方法更具普遍意义,但也有一些只是经验之谈.
>
> 普洛克鲁斯(450 B.C.) 对欧多谟斯(320 B.C.)《(几何学) 史》摘要[1]

公元前379年希密亚斯(Simmias of Thebes)和柏拉图(Plato)在访问埃及的报告(源于普卢塔克(Plutarch of Chaeronea,1 世纪) 的戏剧性描述) 中讲道:"在我们从埃及返回的途中遇到了一批得洛斯人……,他们请求几何学家柏拉图为他们解决一个神在奇怪预言中提出的问题,预言的大意是:得洛斯人和其他希腊人当前面临的种种苦难将会结束,只要他们能够将得洛斯的祭坛体积加倍.得洛斯人不仅没有完全理解预言的意思,而且对于如何建造这样的祭坛感到一筹莫展,于是便请柏拉图帮忙.柏拉图回答道,神嘲笑希腊人疏忽教育,嘲笑我们的无知,他命令我们认真地研究几何,对智力超常又精通于这门学问的人,他们所要做的就是找到两个比例中项,使立方体的各边按比例增加,从而使其体积加倍.后来尼多斯(Cnidus) 的欧多克斯完成了这项工作.然而得洛斯人和希腊人认为神的目的并不在于此,而是想让他们通过学习和讨论数学来平息他们的激情,结束战争和苦难,信奉缪斯女神,建立起一种没有伤害的互利关系."[2]

上面引文中的传奇性描述表明,公元前4世纪之前,希腊人对数学产生了新的看法.他们不仅要计算出问题的数值结果,而且还必须证明结论正确.倍立方体问题 —— 找出一个体积是原来立方体体积二倍的新立方体 —— 与求2的立方根等价,这在数值计算上并不困难,但神注重的是几何作图,而非数值结果,这便需要依赖逻辑论证的几何证明.

约公元前600年,数学的本质发生了改变,这种改变与正在诞生的希腊文明和埃及、巴比伦文明间的巨大差异有关.希腊处于山脉和岛屿的地理环境中,这限制了大规模农业生产的发展,因此

导致的可能结果是希腊人没有建立中央集权政府，希腊的主要行政机构是大小不同的城邦政府(polis)，各个城邦也只有几千人.我们无法断定它们的政府是民主制还是专制制，但一定是按法律统治的，这样就激发其臣民学习辩论的技巧，很可能是这种气氛促进了证明数学的产生.

由于各个城邦离海很近，所以它们与希腊其它的城邦和其它国家有经常的贸易往来，这样，希腊与不同的民族进行交流，希腊人也遍及东地中海的各个地方，同时，希腊不断提高的生活水平吸引了世界各地的人才，他们纷纷来到希腊，这有助于希腊人了解许多基本问题的各种不同的答案后，进一步做出自己的答案.在思想领域的诸多方面，他们不只了解古代流传下了什么问题，还开始问为什么并尽力回答“为什么?”.希腊的思想家们开始逐渐地意识到周围世界是可以被认识的，通过合理的调查研究可以发现世界的特征.于是，他们便热衷于发现和解释物理、生物、医学和政治学等领域中的理论，他们深信，数学是主要的科学之一，是对物质世界所有研究的基础.尽管西方文明认为希腊在文学、艺术和建筑方面没有突出的贡献，但广义地说，作为现代数学的基础和现代技术文明的基石的数学证明思想却归功于希腊.

在本章的开始，我们讨论公元前 6 世纪的希腊数学及其解决问题方法的产生.首先讨论公元前 4 世纪柏拉图和亚里士多德对数学本质的研究及逻辑推理的思想，最后给出写于约公元前 300 年的欧几里得《原本》—— 现存的最早数学著作之一的详细分析，其中也包括对《原本》的内容作出贡献的许多其他早期希腊数学家的工作.

2.1　最早的希腊数学

虽然没有发现早于公元前 300 年的完整的希腊数学著作，但是可以找到许多零散的数学手稿及后来著作中关于早期希腊数学的参考资料.普洛克鲁斯 5 世纪(约在《原本》成书后 800—1000 年)为欧几里得《原本》卷 1 作的评注中包括了对早期希腊数学最完整的资料，这些资料是由欧多谟斯(Eudemus of Rhodes，约 320 B.C.)所撰写的，一般认为，它是对希腊早期数学史的正式综述，可惜其原稿已遗失.这里需提到的希腊最早的数学家是小亚细亚的泰勒斯(Thales，624—547 B.C.)(图 2.1)，许多关于他的故事都是在他去世几百年之后写的，其中包括他在公元前 585 年对日蚀的预测以及应用三角形边角边准则测量海上航船的距离.他发现了等边三角形的两底角相等、两直角相等的定理，证明了圆的直径二等分圆.虽然我们不清楚泰勒斯“证明”的准确方法，但他显然提出了一些逻辑论证.

亚里士多德则提到另一个故事：泰勒斯曾被指责在无用的研究中浪费时间.于是有一次，他用各方面的知识预见到橄榄必然获得特大丰收，便垄断了一地区的榨油机，当橄榄丰收后，种植橄榄者都到他这里来榨油，他因此而获得了巨额财富.他说，如果一个哲学家或数学家想去赚钱，那么事实上他就可以做到.对这些故事，我们不能研究其内容的真实性.



图 2.1　希腊邮票上的泰勒斯.

公元 4 世纪及后期，人们认为泰勒斯是希腊数学传统的开创者，实际上，泰勒斯还是整个希腊科学研究(包括认识到物质现象是受一些不可知规律控制的)的开创者.

2.1.1 毕达哥拉斯及其学派



图 2.2 希腊硬币上的毕达哥拉斯.

毕达哥拉斯(图 2.2) 是希腊的另一位数学家,后人对他的记载较多,他曾在埃及和巴比伦生活了很长时间,大约在公元前 530 年,他被迫离开家乡萨摩斯岛,在意大利半岛南端的克罗托内(Crotone) 定居,在那里他广收门徒,后来形成了著名的毕达哥拉斯学派 —— 一个宗教式哲学学派. 毕达哥拉斯现存的传记都是写于他死后的几个世纪,从这些传记中可知,与其说他是一位理性的思想家,不如说是一位神秘的思想家,而且是一位倍受其追随者尊敬的思想家. 因为没有现存的毕达哥拉斯或其学派的著作,我们只能从包括新毕达哥拉斯学派(neo-Pythagoreans) 的后人的著作中推测其学派的数学学说.

毕达哥拉斯学派的一个重要数学学说是"万物皆数",数,即正整数,形成了宇宙的基本组成原则. 他们不只认为任何事物都具有一个数或可以用数来记,还认为数是所有的物理现象的基础,例如,天空里的一个星座即可用组成它的星的数目刻画,也可用被认为是由一个数代表的几何形式刻画;行星的运动可以根据数的比表示;音调的和谐由数值的比决定:一根拉紧的弦,若取原长的 1/2,可弹出八度音调;若取 2/3,可弹出五度音调;若取 3/4,可弹出四度音调,通过取各种不同的长度,即可以奏出完整的音阶;三边的比是 3 : 4 : 5 的三角形是直角三角形,这样,三角形与数建立了一种联系. 既然毕达哥拉斯学派把数看作是宇宙的基本原则,那么他们对正整数的研究是很自然的事情,而我们今天把正整数称为数论的基本元素.

这一理论的起点是奇数和偶数间的二分法(dichotomy),毕达哥拉斯学派可能用点或小卵石表示数,一个偶数可以用能分成两个相等部分的一行小卵石表示,而一个奇数却不能被分成两个相等的部分. 用小卵石可以很容易地证明一些简单的定理,例如,任意一组偶数的和仍是偶数,而偶数组个奇数的和是偶数,奇数组个奇数的和是奇数(图 2.3).

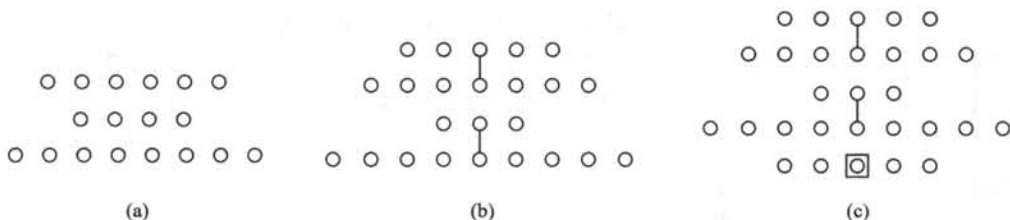
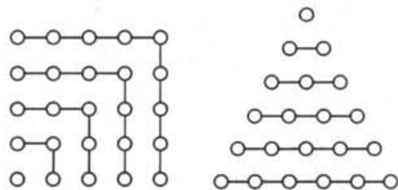


图 2.3 (a) 偶数的和是偶数;(b) 偶数个奇数的和是偶数;(c) 奇数个奇数的和是奇数.

从这些基本概念可得出一些简单推论,如偶数的平方是偶数,奇数的平方是奇数. 数的平方本身也可以用点表示,这样便给出了"形数" 的简单例子. 如果用形数表示一给定数的平方,例如 4 的平方,容易看出,在原来的形数图两边各增加一列(行) 点就可以得出下一个数的平方,增加的点数是 $2\times4+1=9$. 毕达哥拉斯学派把这一结果推广后得出:连续的奇数相加可得到正方形数,例如,$1+3=2^2$,$1+3+5=3^2$,$1+3+5+7=4^2$,这些奇数的 L 形排列一般被称为一个曲尺形排列(图 2.4).


图 2.4 正方形数和三角形数.

另外的形数包括由连续的自然数相加生成的三角形数;具有 $n(n+1)$ 形式的矩形数,它是从 2 开始,连续的奇数相加生成的数(图 2.5),前 4 个矩形数是 2、6、12 和 20,即 1×2、2×3、3×4 和 4×5. 图 2.6 形象地说明了矩形数是三角形数的二倍,正方形是两个连续三角形数的和.

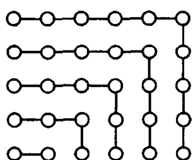


图 2.5 矩形数.

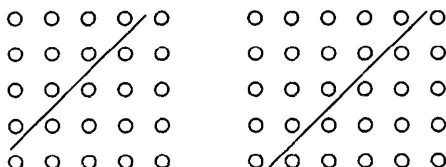


图 2.6 三角形数的两个定理.

关于数论的问题,毕达哥拉斯学派感兴趣的是毕达哥拉斯三元数组(Pythagorean triples) 的构造. 显然他们已经知道,当 n 是奇数,m 是偶数时,

$$\left(n,\frac{n^2-1}{2},\frac{n^2+1}{2}\right)\text{和}\left(m,\left(\frac{m}{2}\right)^2-1,\left(\frac{m}{2}\right)^2+1\right)$$

都是毕达哥拉斯三元数组. 他们是如何根据形数的构造证明第一个结论的呢?他们给出一个注释,即任意一个奇数都是两个连续正方形数的差. 因此,如果奇数本身是一个正方形数, 那么就找到了三个正方形数, 其中一个是另外两个的和 (图 2.7). 为了找到上图中这些正方形数的边,即这个三元数组本身,注意到曲尺形的边已经给出了,因为它是一个奇数的平方,那么较小的正方形数的边可以这样求得,即从曲尺形数中减去 1,然后把剩余的结果二等分;而较小的正方形数的边则是那个较小的边加上 1. 对偶数的情况也可给出类似的证明. 尽管我们没有确切的资料了解毕达哥拉斯学派给出的与三元数组有关的其它结论,但是很可能毕达哥拉斯学派考虑了这些三元数组中的数的奇偶性. 比如,不难证明,在一个毕达哥拉斯三元数组中,如果有一项是奇数,那么另外两项一定是一奇一偶.

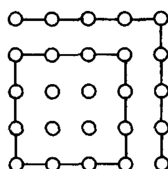


图 2.7 等于两个平方数差的奇平方数.

对毕达哥拉斯三元数组的研究所引出的几何定理 —— 直角三角形中斜边上的正方形等于两直角边上正方形的和,一直被认为是毕达哥拉斯本人发现的,但却没有确凿的证据来证实. 事实上,早在毕达哥拉斯之前,其它的文明已经发现了这一定理. 然而,正是这一在公元前 5 世纪以前就已知道的定理导致了无理数的首次发现.

在毕达哥拉斯学派看来,数字总是与事物的计数联系在一起的,因为计数时要求单位必须保持不变,也就是说每个单位都不能再被分解成其它的单位,尤其是在毕达哥拉斯学派甚至整个希腊的形式数学看来,一个数就意味着"多个单位的组合",也就是说,数是可以计数的. 进一步说,因为 1 并不是多个单位的组合,所以在这种意义下,1 并不像其它正数那样被看作是一个数,甚至亚里士多德也曾说过,2 是最小的"数".

毕达哥拉斯学派把数看作是宇宙的组成基础,他们认为每样东西包括长度都是可度量的. 要度量一个长度,就需要度量单位. 这样,毕达哥拉斯学派假定这样的单元总是可以找到的,在特殊问题中,一旦找到一个度量单元,它就成为一个单位并且不能再分解. 但他们并没有意识到数与量的基本差别,或者说数单位的不可分性与量的量度的无限可分性之间的差别将会带来麻烦.

既然所有的长度都是可度量的,那么毕达哥拉斯学派假定总是可以找到一个度量单元来度量正方形的边和对角线,换句话说,总存在一个长度,而正方形的边和对角线是这个长度的整数倍,遗憾的是,这一假定并不正确. 正方形的边和对角线是不可公度的,两者不存在共同的度量单位. 不管

如何选取度量单位,使得有一个确切的数字能够适合这两条线段之一的长度,另外一条线段总是需要用某个数再加上该单位的一部分才能表示,但是,这个单位又是不可分的.约在公元前430年的这个发现迫使毕达哥拉斯学派放弃他们的万物皆数的基本哲学,并且使得希腊数学家们发展一些新的理论成为可能.

不可公度这一性质到底是如何发现的呢?在亚里士多德的著作中可以找到惟一的线索.他叙述道,如果假定正方形的边和对角线是可公度的,那么可以推出奇数等于偶数的矛盾.不可公度性的一种可能的发现过程如下:假定图2.8中正方形的边 BD 和对角线 DH 是可公度的,即它们都可以表示为其共同的度量单位的倍数.这也可以假定为这两个数字中至少有一个是奇数,若非如此,那么将存在一个更大的共同的度量单位.这样,边上的正方形 $DBHI$ 和对角线上的正方形 $AGFE$ 分别代表一个正方形数,显然 $AGFE$ 是 $DBHI$ 的2倍,所以它表示一个偶正方形数,因此边 $AG = DH$ 也表示一个偶数,并且正方形 $AGFE$ 是四倍的.既然 $BDHI$ 是 $AGEF$ 的一半,那么它一定是二倍正方形,即它表示一个偶正方形,所以说,BD 边一定是偶数,这与 BD、DF 中有一个是奇数相矛盾.因此,这两条线是不可公度的.

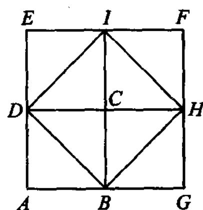


图 2.8 正方形边和对角线的不可公度性(第一种可能).

上面的这一证明说明,证明的思想在那时已扎根于希腊的数学观念中.公元前5世纪的希腊数学是否有了完整的公理体系,是否明确地认识到必须接受某些没有证明的命题,我们没有证据回答这些问题,但他们一定已经认定,在判定特定结论的真实性时,使用逻辑论证的形式是必要的.不可公度性的完整概念的提出是对巴比伦和埃及的数的计算概念的一个突破,在巴比伦的数学方法中,可以给正方形的边和对角线赋一个数值,但希腊数学首次正式地意识到存在着不能找到这个"确切"值的情况.

2.1.2 化圆为方与倍立方体问题

证明的思想和数值计算概念的改变在公元前5世纪中期通过解决两个几何问题的努力而变得更趋明显,希腊数学家对这两个问题 —— 化圆为方(埃及人已开始研究)和倍立方体(如本章开始所述)的研究长达几个世纪.从希腊对这两个问题及稍后的三等分任意角的大量研究使我们知道,希腊数学的中心目的是解决几何问题,从某种程度上说,现存的主要希腊数学著作中的大部分定理是解答这些几何问题的逻辑基础.

希波克拉底(Hippocrates of Chios,公元前5世纪中期,不是希腊名医希波克拉底)是研究倍立方体和化圆为方问题的第一人,希波克拉底可能意识到倍立方体问题与倍边长为 a 的正方形类似,那么可以给 a 和 $2a$ 作一个比例中项 b,即 $a : b = b : 2a$,或 $b^2 = 2a^2$,从而解决问题.从希波克拉底著作的零散记录中可以看出,他非常熟悉比例中的运算.古代的文献记载道,希波克拉底首次提出把倍立方体问题简化为求 a 和 $2a$ 的两个比例中项 b、c.如果 $a : b = b : c = c : 2a$,那么

$$a^3 : b^3 = (a : b)^3 = (a : b)(b : c)(c : 2a) = a : 2a = 1 : 2,$$

且 $b^3 = 2a^3$.但他按这种方法用几何工具并不能作出这两个比例中项,他把这一问题留给了他的后人,在第三章我们对此将讨论.

希波克拉底在化圆为方问题的研究方面也取得了进展,主要是通过证明新月形(由两条圆弧围成的图形)可以化为方形,即可以证明新月形的面积与由直线围成的特定图形的面积相等.要证明新月形可以化为方形,首先他必须证明圆面积的比等于它们直径上正方形面积的比,但我们并不知

道他是如何做的.然而,他可以化四分之一圆为方形(图2.9),设 AB 是以 D 点为圆心的圆的直径,AC、CB 是圆内接正方形的两条边,AEC 是以 AC 为直径的半圆,因为 $AB^2 = 2AC^2$,所以半圆 ACB = 2(半圆 AEC),另外,半圆 ACB = 2(四分之一圆 ADC),所以半圆 AEC = 四分之一圆 ADC,减去它们的公共部分,即可得出新月形 $AECF$ 与三角形 ADC 相等.

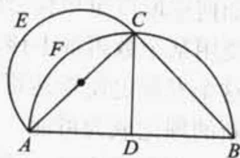


图2.9　希波克拉底的圆内接正方形一边上的新月形.

虽然希波克拉底还给出了把其它新月形或它们的组合图形化为方形的方法,但他仍不能把圆化为方形.他在研究化圆为方和倍立方体问题时建立了大量的几何定理,后来他把这些定理编成了几何基础的第一本著作.

2.2　柏拉图时期

在柏拉图(Plato,429 B.C.—347 B.C.)(图2.10)时期,学者们在倍立方体、化圆为方和不可公度及其对比例理论的影响方面作了重要工作.约公元前385年在雅典建立的柏拉图学苑(Plato's Academy)聚集了当时希腊各地的学者,他们在学术上取得了大的进展,这些学者对数学、哲学和其它领域的知识进行讨论.在此学园建立的700多年后,流传着这样的一个故事,在柏拉图学苑的门口用希腊语刻着 ΑΓΕΩΜΕΤRΗΤΟΣ ΜΗΔΕΙΣ ΕΙΣΙΤΩ 的铭文,大意是,"不懂几何者不得入内",一个不懂几何的学生也是不懂逻辑的,因此他不能理解哲学.

柏拉图的名著《理想国》(The Republic)讨论了胜任治国的哲人国王(philosopher-kings)即国家的理想管理者的教育,其中记述了柏拉图为学苑的学生开设的课程提纲,数学课程包括五个部分,算术(即数的理论)、平面几何、立体几何,天文学和谐音学(音乐).国家的统治者应该"学习算术,但不像商人或店主为了去做生意,而是把它应用于军事或为了灵魂本身去学的,因为这是使灵魂由变化的现实世界转向真理和实在的捷径……,如果是为了知识而非商业目的,这将更有利于我们的教育……,它有强大的力量引导思想前进,推动思想考虑纯粹数的问题,拒绝讨论可看到和触摸到的事物."[3] 也就是说,研究算术是为了训练人的头脑(同时也是为了它的军事应用).柏拉图笔下的算术不只包括上面讨论过的毕达哥拉斯数的理论,还包括其它一些后面我们即将讨论的欧几里得《几何原本》卷7—9中的内容.



图2.10　柏拉图和亚里士多德:拉斐尔名画《雅典学派》.

另外,基于实践的目的,尤其是一个将领在战争中搭建营帐排兵布阵时,需要应用一些几何知识.虽然数学家们在几何学中也讨论诸如平方、加法等实际的操作.但按柏拉图的观点,几何学的目的不是为了实践,而在于获得知识,"而且是永恒存在的知识,而非那种随着时间而变迁的不确定的知识".[4] 因此,在柏拉图看来,几何研究是理论性的,而非实践性的,它和算术一样,也是为了"对灵魂引入真理".比如,柏拉图对人们实际作出的圆与存在于心目中的理想的圆有详细的区分,后者才是几何的研究对象.在实际作图中,不可能同时作出一个圆和与之只有一个公共点的切线,尽管这是数学中的圆与切线的性质.

几何学的下一个研究专题应该是立体几何.柏拉图在《理想国》中抱怨立体几何没有被充分的研究,这是因为"国家认为不值得鼓励人们研究(它)",和因为"缺少指导教师,学生也就不可能做出发现".[5] 但柏拉图认为在这一领域将会有新的发现,实际上从柏拉图的戏剧性对话(约公元前400年)到欧几里得时代,这期间就已经有了新发现,《几何原本》卷6—9中就包括一些立体几何的内容.不管怎样,系统的立体几何知识对以后天文学的研究是必要的,正如柏拉图所说的"天体作圆周

运动",他还认为,作为实体运动着的恒星所表现出的偶然变化和不规则不同于它们运动路径的理想抽象关系和以数字表示的速度及诸如圆这样完美的图形;天文学研究的真正目的是对这种理想体给予数学研究,这可以通过对一些问题进行研究而不必实际遵循每个天体的运动来达到.同样,柏拉图对谐音学中具体的音与抽象出音也作了区分,其实毕达哥拉斯学派已经发现了拉紧的弦所产生的音阶与它们的成正整数比的长度有关.在鼓励那些哲人国王研究谐音学时,柏拉图希望他们超越实际的音乐,即超越具体的弦和声音,以达到"探求哪些数是代表真正的谐音,哪些不是,以及为什么"[6] 这样的抽象水平,也就是说,他们应该像研究天文学和数学一样来研究谐音数学,而不只是关心现实中的乐器或实际的星体.这说明,数学中的比和比例理论是数学在各学科研究中必须应用的主要部分,这些理论包括在欧几里得《几何原本》卷5中.

虽然我们不清楚《理想国》中记述的这些课程在柏拉图学苑中是否真的被讲授,但可以肯定,柏拉图引进了当时最好的数学家来授课和进行研究,其中包括蒂奥泰斯(Theaetetus,约417 B.C.—369 B.C.)和欧多克斯(Eudoxus,约408 B.C.—355 B.C.),而学苑中最杰出的人物是亚里士多德.

2.3 亚里士多德

亚里士多德(Aristotle,384 B.C.—322 B.C.)(图2.11)从18岁开始一直在雅典的柏拉图学苑从事研究,直到公元前347年柏拉图去世.之后不久,应马其顿国王菲利普二世(Philip Ⅱ)的邀请,去当菲利普的儿子亚历山大(Alexander)的老师.亚历山大在公元前335年继承王位后,成功地征服了地中海国家(图2.12),同时亚里士多德返回雅典建立了自己的学派——吕园学派(the Lyceum),并在那里写作、授业和与高才的学生进行讨论,直到去世.亚里士多德的写作涉及许多方面,包括政治、道德、(哲学上的)认知论、物理学和生物学,他在逻辑学方面的工作在数学领域中影响最大.



图2.11 亚里士多德塑像.

2.3.1 逻辑学

在欧几里得之前的数学著作中,虽然只有关于逻辑论证零散的资料(其中一些在前面提到的希波克拉底的著作中出现),但至少在公元前6世纪,希腊思想家们已经开始研究逻辑推理的概念.活跃的国家政治生活鼓励人们开展讨论和发展辩论的技巧,从许多当时的哲学著作中可以找到有关这方面的例子,诸如伊利亚学派的巴门尼德(Parmennides,公元前6世纪后期)及其弟子芝诺(Zeno,公元前5世纪)的著作详细地描述了各种辩论技巧,其中如"归谬法(reductio and absurdum)"——即假定要证明的命题不成立从而引出矛盾;否定后件律(modus tollens)——先证明若A正确,则B也正确,然后证明B不正确,结论是A也不正确.事实上,正是亚里士多德接受了这一已有数百年历史的思想,并首次总结了逻辑论证的基本原则.



图2.12 马背上的亚里山大.

亚里士多德相信,逻辑论证应建立在**三段论**(syllogism)的基础上,"三段论是指由所陈述的事情必定可得出另外的某些结论的论证过程"[7],即三段论是由某些被认为是正确的命题和由此得出的另外一些必定也是正确的命题构成.例如,"猴子是灵长类动物,而灵长类动物是哺乳动物,那么

猴子是哺乳动物”是三段论的一种类型，而“天主教徒是基督教徒，而基督教徒不是穆斯林教徒，那么天主教徒不是穆斯林教徒”是三段论的第二种类型.

亚里士多德在澄清三段论的原则后解释道，三段论推理使我们可以用“旧知识”得出新结论.如果三段论的前提正确，那么结论也必定正确，但是，不是任何知识都可以作为三段论的前提，它必须是从大众普遍接受的事实中推出.亚里士多德对每门特殊学科中的基本原理和大众共知的普遍真理加以区分，把前者称作公设，后者称作公理，他举出“两个等量减去同一量后还剩下两个等量”这一公理作为普遍真理的例子.亚里士多德数作为几何学科的特殊真理的例子如“线和直线的定义”，这意味着他假定了直线是存在的.亚里士多德只是对大多数基本的概念才允许假定被定义对象的存在性，而一般来说，人们要定义一个对象，事实上必须要证明它的存在性，例如，“算术中定义的奇、偶数，平方、立方，以及几何中的不可公度量，…… 然而所有这些性质的存在性必需根据公理或以已知的结论为前提加以证明”[8].亚里士多德也列出了一些论证的基本原理，而早期的思想家们已直觉地应用了这些原理，其中的一个原理是，给定的结论不能既为真又为假；另一个原理是一个结论要么为真，要么为假；除此之外，再没有其它的可能.

在亚里士多德看来，按照他的方法，逻辑论证是获得科学知识惟一确定的途径.虽然也可能有其它的途径，但应用三段论的方法证明一个命题，是我们确定命题结论真假的一种方法.然而，因为并不是任何事物都是可被证明的，所以在采用三段论的方法时要注意前提或公理是正确且众所周知的，正如亚里士多德所言，“那种没有正确前提条件的三段论，并不能导出科学知识，也不具有证明的功能”[9]，也就是说，我们可以选择想用的任何公理，并从中得出结论，但是如果想获得知识，那么就必须以“真实的”公理为前提.我们如何才能确定一条公理是真实的呢?亚里士多德的回答是，这些基本前提是通过归纳得出的，即从我们对大量事例的感知中得出的.自亚里士多德时代以来，这个关于基本公理的“真实性”问题就一直被数学家和哲学家们所讨论.另一方面，亚里士多德通过以公理为前提来获取知识和使用证明来得到新结论的原则已经成为现代数学家们的典范.

虽然亚里士多德强调三段论的应用是逻辑论证的基础，但希腊数学家们显然从未用过三段论，他们使用的是其它的论证形式正如直到今天数学家们还在使用的那样.为什么亚里士多德坚持三段论?这一点并不清楚.公元前 3 世纪的斯多克斯学派(the Stoics)对数学证明中实际应用的基本论证形式进行过详细的分析，其中以克里斯帕斯(Chrysippus, 280 B.C—206 B.C.)的分析最为细致.这种逻辑形式不是以亚里士多德的三段论为基础，而是以命题为基础.克里斯帕斯推理的基本规则可用传统的名词描述如下，若 p、q、r 分别代表命题：

(1) 假言推理(Modus ponens)

如果 p，那么 q；

p，

因此，q.

(2) 否定后件律(Modus tollens)

如果 p，那么 q；

非 p，

因此非 q.

(3) 假言三段论(Hypothetical syllogism)

如果 p，那么 q，

如果 q，那么 r，

因此，如果 p，那么 r.

(4) 选言三段论(Disjunctive syllogism)

要么 p 要么 q,

非 p,

因此 q.

例如,由命题"如果是白天,那么天是亮的"和"现在是白天",根据假言推理就可断言"天是亮的";由命题"如果是白天,那么天是亮的",和"现在天不亮",则根据否定后件律可以断言,现在"不是白天";再由假言三段论,从命题"如果天是亮的,那么我可以看见许多事物"可断言,"如果是白天,那么我可以看见许多事物";最后,根据选言三段论,由命题"要么是白天要么是黑夜"和"现在不是白天",我们可以断言,"现在是黑夜".

2.3.2 数和量的比较

亚里士多德另一项贡献是在数学中区分了数(number)与量(magnitude).毕达哥拉斯学派坚持认为数是万物之本,但亚里士多德拒绝接受这种观点.虽然亚里士多德把数和量放在一个统一的范畴——数量(quantity)里,但他又把这一范畴分为两类,抽象(的数)和连续(的量),作为后者的例子,他提到线、面、体和时间.这两类概念主要的不同之处是"量可以被分成无限的可分的部分"[10],而数的基础是不可再分的单元.因此,量只能由可分的元素组成,而数必然是由不可分元素组成.

亚里士多德进一步阐明了他的"**接续**(in succession)"和"**连续**(continuous)"的概念,事物间没有和自己同类的中间事物被称为是接续的,如3和4是接续的.事物间彼此接触且"它们的每一个接触限都产生一个同类的事物"[11],那么就称它是连续的,有一个公共端点的线段是连续的,因为点必须互相接触且有一个公共的接触限,它们才能构成直线.但由于点没有部分的,这是不可能的.直线上的点也不可能是接续的,否则将会有"下一个点",但直线上两点间的是一条线段,我们总可在线段上找到另外一个点.

现今我们认为线段是由无穷个点组成的集合,但这在亚里士多德看来是毫无意义的,他没有想到过完整或实在的无穷,虽然他也曾用过"无穷"一词,但他只是把它看作是潜在的.比如,我们总可以按要求去二等分一个连续的量,也可以计算出这些二等分量,但我们永不能达到二等分的尽头.进一步说,数学家们确实不需要像无限直线等这样的无限量,他们需要的只是假定例如任意长的直线是存在的.

2.3.3 芝诺悖论

亚里士多德对无穷、不可分、连续和不连续概念的大量讨论的一个原因是他想反驳芝诺(Zeno)的悖论,芝诺提出这些悖论可能是想要说明当时运动的概念不是十分清晰,同时也是为了说明任何对空间和时间的分割方法都会导致问题.

芝诺的第一个悖论,*二分法*,"断言运动不存在,因为运动着的物体要达到终点,首先必须经过路途的一半"[12](当然它必须先走完这一半的一半,依此类推),这里最基本的争论是物体不可能在无限多个时间段内走完有限的路途;第二个悖论是*阿基里斯*,它断言:"比赛中,跑得最快的阿基里斯永远赶不上慢慢爬行的乌龟,因为要追上龟,他必须先到达乌龟的出发点,因此乌龟总是在阿基里斯的前面".[13] 亚里士多德在驳斥这些悖论时承认,时间就像距离一样,是无限可分的,但他并不被物体可以在有限的时间内走过无限多段距离所困扰,因为"如果在有限的时间内一个事物不能与数量上无限的事物相联系,那么它总可与无限可分的事物相联系,在这种意义上,时间本身也是无

限的"[14]. 事实上,在这两个悖论中一旦给定了运动,我们就可以计算出运动的物体何时到达终点,阿基里斯何时超过乌龟.

芝诺的第三和第四个悖论是要说明,当我们说由不可分的元素构成一个连续的量时,将有什么情况发生.飞箭不动悖论说明了"如果处于一定空间的任一物体是静止的,那么运动着的物体在任意时刻总是处于一定的空间,因此飞箭是静止的."[15] 也就是说,如果有不可分的时刻,飞箭在这些时刻静止不动.另外,既然时间只是由时刻组成,那么飞箭就将总是处在静止状态.亚里士多德对这一悖论的反驳是,不仅不存在不可分的时刻,而且运动本身也只能在一个时间段内被定义.而现代对这一悖论是这样反驳的,因为运动是通过极限的观点来定义的,所以可以否定第一个前提.

运动场悖论是说,假设有三个同样物体的集合,A 静止不动,B 经过 A 向右移动,C 以同样的速度向左移动,假设 B 移动到了 A 右面的某个位置,C 移动到了 A 左面的某个位置,且最初在 A_4 下方的 B_1 移到了 A_5 下方,而最初在 A_5 下方的 C_1 移到了 A_4 的下方(图 2.13).芝诺假设物体是空间不可分的单元,它们在不可分的时间单元内移动到了其新位置,但是,既然在某一时刻 B_1 一定在 C_1 的正上方,那么有两种可能出现,要么两个物体没相遇,则物体根本没有运动,要么在不可分的时刻,每个物体各处于两个不同的位置,因此时刻事实上是可分的.亚里士多德相信,他已经驳倒了这个悖论,因为他已经否定了最初的假定,即时间是由不可分的时刻组成.

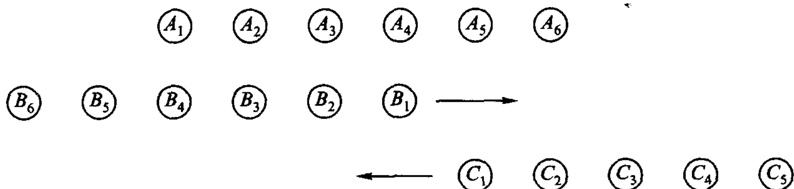


图 2.13 芝诺运动场悖论.

这些悖论的论战贯穿了整个历史,而芝诺悖论的思想及亚里士多德的反驳引起诸多学者深思,甚至现代数学家们在处理无穷或无穷小的概念时也必须对其假定认真斟酌.在希腊时期,悖论及对悖论的反驳是促使对连续量与离散数作出区分的一个重要因素,这对于亚里士多德乃至欧几里得来说是多么的重要.

2.4 欧几里得与《原本》

希腊时期乃至整个人类历史上最重要的数学著作就是欧几里得的《原本》,它写成于约 2300 年前,是除《圣经》外发行量最大的著作.《原本》已经几乎被译成了每种语言,自从它首次印刷以来,便不断地被各个国家再次印刷出版.现在的读者可能难以理解这一著作,没有例子,没有情景,没有注释,也没有计算,而只有简单的定义、公理、定理和证明.但是,这一著作已经被深入细致地研究过.许多著名数学家的传记中都指出,欧几里得这本书是最早把他们引入数学研究并激励和促使他们成为数学家的著作.《原本》给他们提供了"纯数学"的模式,严密的公理、准确的定义、仔细陈述的定理和逻辑一致的证明.虽然此前可能见过类似的著作,但欧几里得的《原本》是惟一能够幸存至今的,这可能因为它是在柏拉图学派发展了比例理论和无理数理论基础及亚里士多德细致地对数和量进行区分之后写成的第一部数学著作,因此它的内容"完整",结构严密,因为当时数学团体规模不大,一旦他们意识到《原本》的优越性,那么就没必要在课程中保留并使用其它相形见细的

著作.通过下面对《原本》详细的分析,不仅可以使读者深入了解它的优越之处,而且还能对这部著作所包括的希腊数学各个部分的起源作出解释,这里所讨论的每一方面都体现出《原本》对后来数学发展的重要性.

我们对《原本》作者的生平了解甚少,在他生后约 750 年的普洛克鲁斯(Proclus,410—485) 的《评注》中有少量关于欧几里得的记载:

> 欧几里得是在这些人(科洛丰的赫莫梯姆斯和门德的菲力浦斯,均为柏拉图的学生)之后不久与他的《原本》一同出名的,在《原本》中,他把欧多克斯的许多定理系统化,完善了蒂奥泰德(Theaetetus) 的许多定理,把前人十分松散的命题建立在不可反驳的证明形式上.他生活在托勒密一世的时代,这是因为托勒密一世晚期的阿基米德曾提到欧几里得.还有一种说法是,托勒密曾问欧几里得,除了《原本》之外,有没有其它学习几何的捷径,因此他一定晚于柏拉图学派而早于埃拉托塞尼(Eratosthenes) 和阿基米德.[16]

一般认为,欧几里得曾在亚里山大博物院(Museum) 和图书馆任教并进行学术活动(图 2.14),这个复合机构是在约公元前 300 年,托勒密一世索特(Soter,马其顿的将军,在公元前 323 年亚里山大死后,统治了埃及)建立的."Museum"意思是"缪斯女神庙",它实际上是政府研究组织,它为成员提供生活费和饮食费用且免征个人所得税,托勒密一世及其继承者希望用这种方法从希腊各地吸引优秀人才.事实上,博物院和图书馆不久便成为希腊人文科学和自然科学最高级的学术研究中心,因为在那里也聚集了许多年轻的学生,所以这些成员不久便转向教学.图书馆的目标则是搜集希腊所有最实用的文献并使之系统化,为此,外出船队被指示在他们返航时从经过的每个港口带回各种书卷.有一则故事讲道,公元前 247—221 年,托勒密三世在位,他从雅典借来剧作家埃斯库罗斯(Aeschylus)、索福克勒斯(Sophocles) 和欧里庇得斯(Euripides) 的剧本原版收藏起来,而把这些剧本的复印本还给他们.图书馆最后共搜集了 500 000 多册涉及各个知识领域的书籍,虽然有一些在多次战乱中遭到破坏,但直到 4 世纪,其中仍有许多保存完好无损.



图 2.14 欧几里得(源自拉斐尔名画——雅典学派).

约在 2300 年前,欧几里得已写就了《原本》,但我们却没有找到一本当时的原著.在埃及发现的公元前 225 年的陶瓷碎片上有现存最早的《原本》片断,上面写有《原本》卷 8 中的两个命题.约公元前 100 年的一份纸草书上包括有《原本》卷 2 中的部分内容.后来许多人对《原本》作了注释和评论或加入新的引理,其中尤为重要的是亚里山大的塞翁(Theon,公元 4 世纪) 的修订本,现存多数的欧几里得《原本》版本都是以这一修订本为底本.然而,现存最早的版本是收藏在牛津大学鲍德莱(Bodleian) 图书馆的一份公元 888 年的手稿.但在梵蒂冈(Vitacan) 图书馆收藏着 10 世纪的一个手稿,这一手稿不是以塞翁之前的本子为底本的.丹麦学者海伯格(J.L.Heiberg)19 世纪 80 年代把这一手稿与其它几个以塞翁的为底本的手稿作了细致的比较,编辑了尽可能接近原著的最权威的希腊文本(海伯格对其它几本重要的著作也作了同样的工作).这里所讨论的内容都是节选自希思(Thomas Heath) 的海伯格希腊文本的英文译本.

欧几里得的《原本》共由 13 卷组成,但内容在整体上并不很统一,从其内部结构和援引的希腊数学史的资料可以看出,《原本》是欧几里得把当时许多数学著作中的不同内容重新组织后写的一部概要.但是,《原本》在结构上却是完整的,著作的前六卷对二维几何量作了比较完整的叙述,其中包括亚里士多德对数和量的基本区分,卷 7—9 叙述了数的理论,实际上,欧几里得把比例理论分

为两个独立的部分进行讨论;卷5和7分别讨论量和数;卷10中引入公度量和不可公度量的概念,指出可公度量在比例中可以把它们看作数来处理,同时对一些不可公度量进行了分类,从而给出了两个概念间的联系;卷11和12讨论了三维立体几何;卷13中构作了五种正多边形并按卷10的体系对一些直线进行分类.

需说明的是,除算术计算的具体方法外,《原本》的内容以不同的形式几乎包括了所有我们在第一章中提到的古代数学,但它的方法与古代截然不同.早期文明的数学总是包括数字和测量,把解决问题的数值算法看得特别重要,而欧几里得的数学与算术完全不同,他的著作中不包括测量,也只涉及了少量的正整数,同时也没有用到肘尺(cubit)、亩(acre)和角的度数,而对角的测量标准只是直角.

我们不免要问,欧几里得所用的与早期文明的思想有关的资料是来自希腊本土的原创,还根据是这些早期文明的知识改编而成?本章将给出一些例证,但仍不能得出此问题的具体答案.

2.4.1　定义和公设

正如亚里士多德所言,科学著作必须从定义和公理开始,因此欧几里得《原本》的很多卷的开始都对要讨论的对象作了定义.下面介绍《原本》卷1中的定义.

补遗 2.1	欧几里得《原本》卷1中的定义
	1.点是没有部分的. 2.线是没有宽度的长. 3.线的两端是点. 4.直线是它上面均匀分布着点的线. 5.面是只有长度和宽度的. 6.面的边界是线. 7.平面是它上面均匀分布着线的面. 8.平面角是在同一平面内但不在同一条直线上的两条相交线相互的倾斜. 9.当包含角的两条线是在同一条直线时,这个角叫做平角. 10.当一条直线和另一条直线交成的邻角彼此相等时,这些角的每一个叫做直角,而且称一条直线垂直于另一条直线. 15.圆是由一条线包围着的平面图形,其内有一点与这条线上的所有点连成的线段都相等. 16.而且把这个点叫做圆心. 17.圆的直径是过圆心而两个端点在圆周上的任意线段,且把圆二等分. 18.半圆是直径和由它截得的圆弧所围成的图形,而且半圆的心和圆心相同. 23.平行直线是在同一平面内的直线,它们向两个方向无限延长,不论在哪个方向都不相交.

按现代定义的标准,欧几里得的前几个定义并无实际意义,根据对对象的理解不可能对它们都作出定义,尤其在今天,我们对诸如点或直线等不作定义.承袭亚里士多德的传统,欧几里得不只用这些定义解释特定的词汇,而且说明所定义的对象是存在的,因此,他对点、线、直线、面、平面和平面角的定义不仅有助于我们理解这些概念,还使得我们能够确定这些对象是存在的.另外,定义3和6分别相对于点和线来定义线和面.

由定义9我们知道,欧几里得所说的线并不一定指直线,所以一个角也可以由圆弧构成,他在

卷 1 中给了平角的定义. 同样, 在定义 15 中他用线表示圆周. 定义 17 中讲到了泰勒斯的定理, 即圆的直径二等分圆, 在这一定理基础上, 欧几里得明确定义了半圆. 其它的定义相对来说要简单一些, 其中有许多在欧几里得之前的著作中已有描述, 但欧几里得对平行线的定义有些让人感到奇怪, 在实际中如何应用这一定义, 又如何确定两条直线永不相交呢? 因为不可能作出无限长的直线, 所以欧几里得承袭了柏拉图式的理想平面上理想直线的概念.

亚里士多德曾说, 我们必须接受某些命题是正确的, 欧几里得把这些命题分为两类, 第一类是几何学中特有的真理 —— 公设:

1. 从任意一点到任意一点可作一直线.
2. 一条有限直线可以不断延长.
3. 以任意中心和任意的距离可以画圆.
4. 凡直角都彼此相等.
5. 若一直线落在两直线上所构成的同旁内角和小于两直角, 那么把两直线无限延长, 它们将在同旁内角和小于两直角的一侧相交.

《原本》中有大量的命题是要求作出满足特定性质的图形, 而前三个公设是构作这些图形的基础.《原本》中的许多命题实际上是希腊数学家们欲解决的某类问题的实例, 正是对这些问题答案的寻求导致了许多定理的发现, 而这些定理又是许多命题成立的基础. 欧几里得在作满足特定条件的线、圆或特殊类型的图形时, 如定义 20 中的等边三角形, 总是以前三个公设为基础说明怎样作出线、圆或特殊图形.

众所周知, 欧几里得作图是以直尺和圆规为工具的, 公设 1 和 2 表明可以用直尺在两点间作直线或把直线延长, 而公设 3 表明已知半径, 以任一已知点为圆心可以画圆. 那么为什么欧几里得的作图工具只限于直尺和圆规, 而他之前或之后的数学家却用不同类型的工具去解决问题呢? 欧几里得没有给出确定的答案, 而是指出, 这些作图是他必须去研究的基本的结论, 即是"初等(或基本)"的, 而其它的作图则是属于"高等"数学.

公设 4 和 5 比公设 3 更晦涩难懂. 无庸置疑, 公设 4 是成立的, 因为欧几里得把直角作为角的测量标准, 所以公设 4 对理解公设 5 的意义是必要的.

公设 5, 即所谓的"平行公设", 是欧几里得的 5 个公设中最复杂的一个, 它不像前四个那么不证自明. 这一公设是说, 如果直线 l 与直线 m 和 n 相交后所成角 1 与 2 的和小于两直角的和, 那么直线 m 和 n 最终要在 A 的方向上相交(图 2.15).

不管前 4 个公设的历史如何, 事实上亚里士多德曾指出, 在他的时代平行理论并没有坚实的基础, 这使我们相信欧几里得所提出的第 5 公设是平行理论的起点. 然而, 当欧几里得提出这一公设后, 当时的许多数学家就试图证明它是一个定理而不是公设, 但所有的这种尝试都以失败告终, 而欧几里得把它作为一条公设, 从这一点上我们可以看出他的数学天才. 另一方面, 现代数学家已经发现欧几里得还应用了一些其它的他并未陈述的公设, 在后面我们将提及这些公设. 总体上看, 这些未被陈述的公设并不是特别重要. 一般来说,《原本》的逻辑结构为数学建立了一个模式.

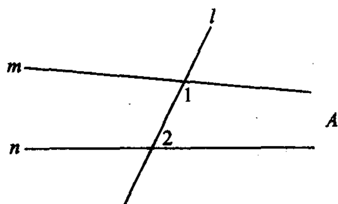


图 2.15　欧几里得《原本》第 5 公设, 平行公设.

在公设之后, 欧几里得总结出了对所有学科都成立的真理, 他把它们称为"公理(common notions)":

1. 等于同量的量彼此相等.

2. 等量加等量,和相等.

3. 等量减等量,差相等.

4. 彼此能重合的图形是全等的.

5. 整体大于部分.

这些公理是不证自明的,其中的前三个在初等几何中经常使用.欧几里得在许多几何问题中应用了这五个公理.

2.4.2 基本命题

《原本》卷1中的许多定理是我们熟悉的,如果不考虑定理的证明,那么其中多数定理的结论可以追溯到希腊几何的最早时期,欧几里得的工作也可能只是组织了卷1的内容和总结了毕达哥拉斯定理.

卷1的前3个问题是作图问题.命题1是讲在一条已知线段上作一个等边三角形(图2.16).设 *AB* 是已知的线段,以点 *A* 为圆心、线段 *AB* 为半径作圆 *BCD*,再以点 *B* 为圆心、*BA* 为半径作圆 *ACE*,连接 *A*、*B* 和两圆的一个交点 *C* 后得到三角形 *ABC*,可证明所作出的三角形是等边三角形.欧几里得的证明如下,因为 *A* 是圆 *CDB* 的圆心,所以 *AC* 等于 *AB*,又 *B* 是圆 *CAE* 的圆心,所以 *BC* 等于 *AB*,既然 *AC* 和 *BC* 都等于 *AB*,那么根据等于同一个量的量彼此相等可知 *AC*、*AB* 和 *BC* 三者相等,所以三角形 *ABC* 是等边三角形.

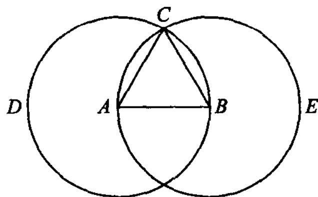


图2.16 《原本》卷1命题1.

因为这是《原本》中的第一个命题,所以欧几里得的证明中只有定义、公设和公理,他根据公设3作两个圆,根据公设2作两条线段,根据圆的定义(定义15)得出 *AC* 等于 *AB*,*BC* 等于 *BA*,最后由公理1可得三角形的三条边相等.如果我们分析一下欧几里得的这一证明过程,他并没有应用亚里士多德的三段论方法,而是采用假言推理的形式给出命题的逻辑结构.

许多注释者指出这个证明中存在逻辑问题,欧几里得怎么知道圆 *BCD* 和 *ACE* 相交呢?因此必须要用到连续性公设,也可能是因为这样的公设太显然了,所以欧几里得才没有给出,而连续性公设直到19世纪才提出,在第17章中我们将对此予以讨论.

命题2和3是作与已知线段相等的线段,命题4是关于一个特殊几何作图的结论,它是本卷的第一个定理,也是三角形全等的第一个定理,通常被简称为边角边定理或 *SAS*.

命题 Ⅰ-4 如果有一个三角形的两边分别与另一个三角形的两边相等,且相等线段所夹的角对应相等,那么这两个三角形全等.

这里所用的词"全等"是欧几里得结论的现代简称,他的结论是一个三角形的每条边与另一三角形的边对应相等.欧几里得不但证明了这一定理,同时他也用重叠的方法证明了两三角形全等的第二个定理,即边边边定理或 *SSS*(命题8),也就是说,他假设把第一个三角形从原来的位置移到第二个三角形上,把一边放在与其对应相等的边上,对应的角也重合.欧几里得默认这种保持形状不变的移动是可行的.而19世纪的数学家们试图把这一定理本身作为一个公设,为了后面的参考,这里我们给出三角形全等的第三个定理,角边角定理或 *ASA*,其前提是一个三角形的两个角和一边与另一个三角形的对应相等,此定理的证明与命题26相同.

命题 Ⅰ-5 等腰三角形中,两底角彼此相等,若向下延长两腰,则三角形底边下方的两个角也相等.

欧几里得的证明如下：设等腰三角形 ABC 中，$AB = AC$（图2.17），AD 和 AE 是两等边的延长线，在 BD 上任取一点 F，再作 AG 等于 AF（命题3），连接线段 CF 和 GB. 接着根据命题4证明三角形 AFC 与三角形 AGB 全等. 因为 $AF = AG$，$AB = AC$，且它们的夹角相同，再由命题4可知三角形 BFC 与 CGB 全等，所以对应角 FBC 与 GCB 相等，又因为角 BCF 等于角 CBG，角 ACF 等于角 ABG，欧几里得由公理3推出等腰三角形的两个底角 ACB 和 ABC 相等.

我们可以用假言三段论分析欧几里得的这一证明结构. 如果 p 代表"ABC 是等腰三角形"，q 代表"底角相等"，那么要证明的定理即是"如果 p，那么 q". 但是如果 p_1 代表"三角形 AFC 与三角形 AGB 全等"，p_2 代表"三角形 BFC 与三角形 CGB 全等"，那么欧几里得已经证明了"如果 p，那么 p_1"，"如果 p_1，那么 p_2"，"如果 p_2，那么 q". 在两次应用假言三段论规律后，欧几里得得出"如果 p，那么 q". 像这样连环式的证明是欧几里得证明中的典型，在我们对其它命题的研究中将不再作这样的分析.

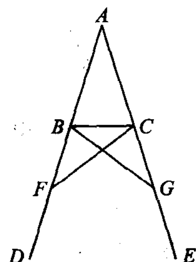


图2.17 《原本》卷1命题5.

当我们分析欧几里得种种论证的逻辑结构时，应当记住这些论证的有效性并不能使我们明白欧几里得或其继承者是如何开始他们的论证的. 欧几里得本人也没有说明为什么他以某种特定的思路进行论证. 因为对结果"为什么"如此从来没有交代，因此把他的著作当作教材就有些困难了. 欧几里得及其继承者很可能在亚历山大和其它学校给学生讲授如何分析问题和给出证明. 但在古代文献中关于分析方法惟一详细的讨论是由公元3世纪的帕普斯(Pappus)给出，这种方法先假定已知的定理是正确的或要作的图形已被作出，然后进行一系列推导，直到得出一个已经被确定是正确的命题为止. 分析法是发现证明的一个特别重要的手段，在第5章中我们将进一步讨论这一方法.

然而欧几里得在运用归谬法时却明确使用过某种形式的分析. 归谬法即假定提出的定理 p 不正确，或非 p 正确，然后推导一系列结论直到得出一个矛盾. 而这个矛盾一般以三种形式之一出现：或 p 正确（这与非 p 矛盾）；或对另外的命题 q，可得出 q 与非 q 矛盾；或得出与某个正确的命题 r 相矛盾的非 r，因此可以推出要证的结论，即 p 正确.

欧几里得用归谬法证明的第一个例子是《原本》卷1的命题6.

命题 Ⅰ－6 如果一个三角形中的两个角相等，那么等角所对应的边也相等.

已知三角形 ABC 中，$\angle B = \angle C$，欧几里得假设 AB 不等于 AC（图2.18），那么将有两种情况出现：$AB > AC$ 或 $AB < AC$. 他首先考虑第一种情况，在 AB 上取一点 D，满足 $DB = AC$，连接 DC，因为 $\angle DBC = \angle ACB$，所以由 SAS 定理可得到 ΔBDC 与 ΔCAB 全等，即部分等于整体，这与公理5矛盾. 类似地可以证明"$AB < AC$"也是错误的，所以 $AB = AC$. 这种归谬法是欧几里得证明两条线段或两个角相等的典型方法，对这样的证明方法，他虽没有明确提到，但他一般采用三分律(trichotomy principle)，即对两个量 a 和 b，$a = b$，$a > b$ 或 $a < b$ 三个条件中总有一个成立，换言之，他假定命题"非($a = b$)"与"$a > b$ 或 $a < b$"等价，那么他就可得到上面要证的结果.

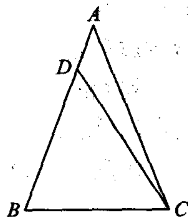


图2.18 《原本》卷1命题6.

命题9到12给出了一些必要的作图问题. 欧几里得讨论了二等分一个直线角、二等分一条直线段、过直线上一点作直线的垂线和过直线外一点作直线的垂线的方法. 命题13是说：一条直线与另一条直线相交所成的同旁角或者是两个直角，或者它们的和等于两个直角. 本卷后面许多命题是以该命题为基础的. 命题14是命题13的逆命题，而命题15是如果两直线相交，那么它们所成的对顶角相等.

命题 16 的证明有些蹊跷，因为欧几里得在此命题的证明中再次用到了前面未加陈述的假定.

命题 Ⅰ－16 在任意三角形中，如果延长一边，那么其外角大于任何一个内对角.

延长三角形 ABC 的边 BC 到点 D(图 2.19)，E 是 AC 的中点，连接 BE，欧几里得然后把 BE 延长到点 F，使得 $EF = BE$. 但是，他把线段延长到任意长度的这种做法并没有公设的支持，当然，如果承认这种做法，那么证明便很简单了. 连接 FC 后证明 ΔABE 与 ΔCFE 全等，所以 $\angle BAE = \angle ECF$，又因为 $\angle ECF$ 是 ΔABC 的外角 $\angle ACD$ 的一部分，所以 $\angle ACD > \angle BAE$.

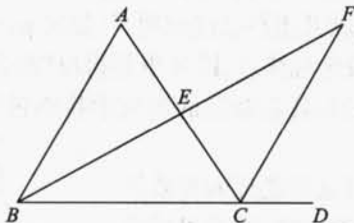


图 2.19 《原本》卷 1 命题 16.

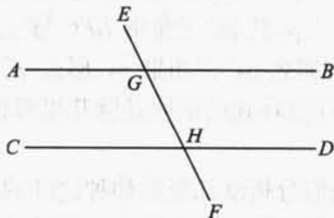


图 2.20 《原本》卷 1 命题 29.

命题 17 是命题 16 的直接推论，三角形的任意两内角的和总是小于两直角. 这一命题在后来导致非欧几何的发现中占有重要地位，我们在第 14 章中将予以详细讨论.

在关于三角形性质的一些命题后，欧几里得从命题 27 开始讨论平行线这一重要概念，在随后的一些命题中他才开始应用有争议的公设 5，事实上，《原本》卷 1 中命题 28 之后的每个命题都是建立在这一公设基础上的.

命题 Ⅰ－29 一条直线与两条平行直线相交，则所成的内错角相等，同位角相等，同旁内角的和等于两直角和.

在命题中，后两个结论可由第一个结论推出，欧几里得应用反证法证明内错角相等. 假设 $\angle AGH > \angle GHD$(图 2.20)，那么 $\angle AGH$ 与 $\angle BGH$ 的和大于 $\angle GHD$ 与 $\angle BGH$ 的和，由命题 13 可知前两个角的和等于两直角和，因此后两个角的和一定小于两直角和，再由公理 5，那么直线 AB 和 CD 最终要相交，但这与它们平行相矛盾. 如果假设 $\angle AGH < \angle GHD$，同样也可得出矛盾. 故命题得证.

以命题 29 为基础的其它结论中，有一个定理是三角形的三个内角和等于两直角和，这一定理与命题 17 一样，在非欧几何的发现中也占有重要地位.

命题 33 研究了平行四边形. 虽然欧几里得没有研究测量或度量公式，但他也证明了一些有关平行四边形和三角形面积比较的结论. 例如，他证明了在同底上且在相同两平行线间(即等高) 的平行四边形彼此相等(命题 36)，一个平行四边形是与它同在两平行线间且同底的一个三角形的两倍(命题 41).

卷 1 的最后给出了可能是《原本》中最著名的定理，即毕达哥拉斯定理及其逆定理(图 2.21).



图 2.21 希腊邮票上的毕达哥拉斯定理.

命题 Ⅰ－47 直角三角形中，斜边上的正方形等于两直角边上的正方形的和.

欧几里得通过作图来证明这一命题. 过三角形的直角顶点 A 作与斜边上的正方形的边 DE 垂直的线段，再证明矩形 BL 等于直角边 AB 上的正方形，矩

形 CL 等于 AC 上的正方形(图 2.22). 此处欧几里得很可能承袭了希腊早期用相似观念证明的方法. 事实上, 由三角形 ABN、CAN 和 CBA 相似可知, AB 上的正方形等于以 BC、BN 为边的矩形(即矩形 BL), 而 AC 上的正方形等于以 BC、NC 为边的矩形(即矩形 CL), 因为两个矩形的和等于 BC 上的正方形, 所以定理得证.

欧几里得想尽可能地把这一定理放在《原本》中的前面, 但有关相似的命题直到卷 5 和卷 6 才出现, 所以他不得不使用另外的方法证明直角边上的两正方形等于斜边上的正方形, 为此他应用了命题 41 的结论, 一个平行四边形是与它同在两平行线间且同底的一个三角形的两倍, 即有矩形 BL 是 ΔABD 的两倍, 正方形 AF 是 ΔFBC 的两倍, 而由 SAS 定理可知这两个三角形是全等的. 对其它的三角形和矩形也有类似的结论, 从而证明了这一定理.

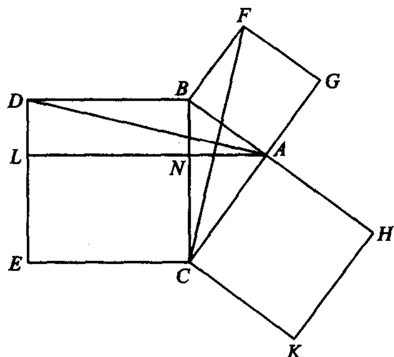


图 2.22　欧几里得《原本》中的毕达哥拉斯定理.

2.4.3　几何代数

《原本》中卷 2 与卷 1 的风格截然不同, 卷 2 主要讨论了不同的矩形与正方形的关系, 其中多数都可用现代的代数符号解释. 实际上, 卷 2 中的命题与卷 1 的命题 43—45、卷 6 的命题 27—30 可以组成所谓的"几何代数", 即用几何图形表示代数概念和运算. 近些年来, 对希腊数学感兴趣的学者们的主要争论的话题便是几何代数, 究竟希腊数学家对代数是否感兴趣, 希腊数学的这方面内容是否直接或间接地从巴比伦借鉴而来? 他们围绕这些问题展开讨论. 另一方面, 作为几何知识的一部分, 这些命题的确形成了一个相对完整的结构, 而把这些命题转换成简单的代数规则或巴比伦解决二次方程的标准步骤也是很容易的.[17]

在卷 2 中, 欧几里得从一个定义开始:

任何矩形都是由形成直角的两条线段**构成**的.

这一定义说明欧几里得对几何的应用. 这一定义并没有说矩形的面积等于其长和宽的乘积, 因为欧几里得无法给出任意的长度乘法的定义, 所以他从未把长和宽相乘. 有些命题中, 他用数字(即正整数) 乘以长度, 或者是只提到由两条线所包含的矩形, 那么便产生一个疑问, 是否应该把欧几里得的"矩形" 只是简单地理解为一个"乘积"?

卷 2 命题 1 便是欧几里得对上面的定义应用的一个例子:

命题 Ⅱ - 1　如果有两条直线, 其中一条被截成任意几段, 那么原来两条线段构成的矩形等于各个小段和未截的那条线段构成的矩形之和.

作为一个代数结果, 这一命题是说, 已知长为 l 和 w 的两条线段, 且 w 被分成几个部分, 如 $w = a + b + c$, 通过这些线段可以确定矩形的面积 lw 等于几个小矩形的面积和, 即 $la + lb + lc$(图 2.23). 换言之, 定理是讲分配律: $l(a + b + c) = la + lb + lc$, 奇怪的是这一结果在证明毕达哥拉斯定理中证明 BC 上的正方形等于矩形 BL 和 CL 的面积时已经使用了, 如果欧几里得在卷 1 中觉得这一几何结果是显然的, 那么他又为什么在后面给出了完整的证明? 除非可能是他把卷 2 作为一些代数结果的几何转换.

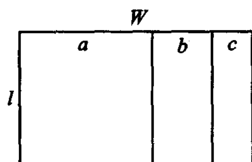


图 2.23　《原本》卷 2 命题 1: $l(a + b + c) = la + lb + lc$.

命题 Ⅱ-4 如果任意两分一条线段,那么在整条线段上的正方形等于各个小线段上的正方形的和再加上由两条小线段构成的矩形的二倍.

这一命题可以简单地表示成代数中的二项式等式:$(a+b)^2=a^2+b^2+2ab$,它是我们在第一章中讨论的求平方根算法的基础(图2.24).欧几里得对这一命题的证明相当复杂,因为他必须证明整个图形中的每一部分是正方形和矩形.现代的证明是把此结论化简为命题1的形式.那么,这一结论究竟应该被看作为几何的还是代数的呢?

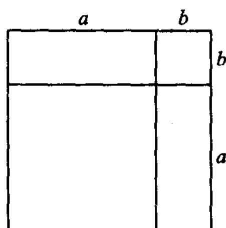


图2.24 《原本》卷2命题4:$(a+b)^2=a^2+b^2+2ab$.

接下来的两个命题可以转换成二次方程标准代数解的几何解释.

命题 Ⅱ-5 如果把线段分别分成相等和不相等的两部分,那么由这不等的两部分构成的矩形与原线段的两个分点间的线段上的正方形的和等于原线段一半上的正方形.

命题 Ⅱ-6 如果二等分一条线段,并在同一线段上再加上另一线段,那么合成的线段与所加的线段构成的矩形与原线段一半上的正方形的和等于原线段的一半加上另一线段后所成的线段上的正方形.

图2.25可以让我们清楚理解这两个命题的意思.在每个图中,AB记为b,AC和BC记为$b/2$,BD记为x,那么命题5可以变成$(b-x)x+(b/2-x)^2=(b/2)^2$,命题6即是$(b+x)x+(b/2)^2=(b/2+x)^2$.将第一个等式写成$(b/2-x)^2=(b/2)^2-c$,可得出二次方程$bx-x^2=c$[或$(b-x)x=c$]的解是

$$x=\frac{b}{2}-\sqrt{\left(\frac{b}{2}\right)^2-c}.$$

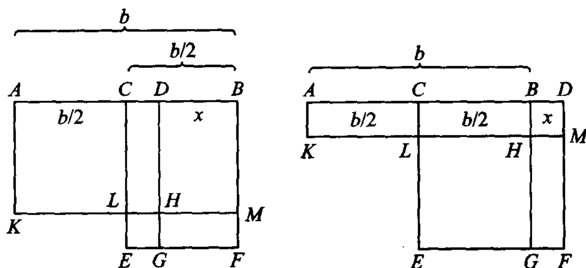


图2.25 《原本》卷2命题5、6.

同样,方程$bx+x^2=c$[或$(b+x)x=c$]也可由第二个等式用类似的公式解出.我们也可以在这两个图中记AD为y,DB为x,那么第一个结果可以变成标准的巴比伦形式$x+y=b,xy=c$,第二个结果变成$y-x=b,yx=c$.注意图2.25与图1.27基本相似,而图1.27则是代表了巴比伦书记员解这第一组方程的可能方法.

类似地,考虑命题9:

命题 Ⅱ-9 如果把一条线段分成相等和不等的两部分,那么不等的线段上的正方形的和等于原线段一半上的正方形与两个分点之间的线段上的正方形的和的二倍.

此命题的一种可能的代数解释是

$$x^2+y^2=2\left(\frac{x+y}{2}\right)^2+2\left(\frac{x-y}{2}\right)^2,$$

这即是巴比伦用来解决 $x - y = b, x^2 + y^2 = c$ 的结果.

当然,欧几里得本人并没有作任何这样的转换,他只是用图 2.25 来证明命题 5 和 6 中有关正方形和矩形的相等性,并用一个包括毕达哥拉斯定理在内的证明来证命题 9(和相似的命题 10).他在整个著作中对于这些命题对解我们所谓的二次方程的应用并没有作任何说明.

在欧几里得看来,这些命题的意义是什么?因为命题 9 在《原本》后面的内容里并没有被使用,我们很难回答这一问题,可能这些命题确是用来解决巴比伦方程组的.但命题 11 的证明中使用了命题 6.

命题 Ⅱ - 11 分已知线段,使得它与所分出的小线段构成的矩形等于另一小线段上的正方形.

这一命题的目的是在线段上找一点 H,使得 $AB \times HB$ 等于 AH 上的正方形(图 2.26).若把这一问题变为代数形式,可设 $AB = a, AH = x$,那么 $HB = a - x$,问题即是解方程

$$a(a - x) = x^2 \quad 或 \quad x^2 + ax = a^2.$$

巴比伦的解答是

$$x = \sqrt{\left(\frac{a}{2}\right)^2 + a^2} - \frac{a}{2}.$$

欧几里得的证明看上去就是这一公式.要得到两数平方和的平方根,较明显的方法是应用已知两直角边是两个平方根的直角三角形的斜边,如上式中即以 a 和 $a/2$ 作为直角边.因此,欧几里得在 AB 上作正方形,再二等分 AC,那么 EB 即为要求的斜边.要从 EB 中减去 $a/2$,他作 $EF = EB$,再从 EF 中减去 AE 得 AF,那么 AF 即是需要的值 x.因为他想在 AB 上取这一长度,故在 AB 上取点 H,使得 $AH = AF$,然后他应用命题 6 证明了点 H 满足条件.

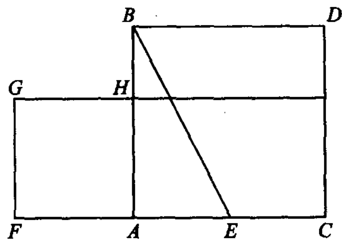


图 2.26 《原本》卷 2 命题 11.

欧几里得显然以几何的形式解出了一个二次方程,有趣的是,他在卷 6 的命题 30 中又解了同样的问题,在那个命题中,他想把一线段分成“中外比(extreme and mean ratio)”,即在已知直线 AB 上找一点 H,使得 $AB : AH = AH : HB$,其代数形式可表示成比例式 $a : x = x(a - x)$,此式可化简为上面的方程,$a : x$ 即是我们熟知的**黄金分割比**,即 $(\sqrt{5} + 1) : 2$,关于黄金比的重要性的研究一直从希腊时期持续到现在.[18]

在讨论卷 2 命题 15 前,我们有必要再看一下卷 1 中的一些内容.

命题 Ⅰ - 44 用已知线段和已知直线角作一个平行四边形,使得它等于已知的三角形.

此命题是已知一个角和等于已知线段的一条边,作一个面积已知的平行四边形,也就是说,把平行四边形“贴合(applied)”到已知的线段上.根据一些文献,面积“贴合”的概念源自毕达哥拉斯学派.如果已知角是直角,那么可以容易地给出这一命题的代数形式.设三角形的面积为 c,已知的线段长为 a,那么问题的目标是确定一条线段 b,使得长和宽分别为 a 和 b 的矩形的面积等于 c,也就是解方程 $ax = c$.既然欧几里得没有处理过量的“分割”,对他来说,要求的结果实际上是解比例式 $c:x = a:1$,其中 1 是已知的单位长度.但是,既然他在卷 1 中没有应用比例理论,那么他不得不使用涉及到面积计算的更复杂的方法.

在卷 1 命题 45 中,欧几里得通过把一个图形简单地划分成若干三角形后,再应用卷 1 命题 44 给出了作一个与已知直线形相等的矩形的方法,他在得出卷 2 命题 14 的结论过程中首先应用了这一命题.

命题 Ⅱ－14　作一个等于已知直线形的正方形.

若用代数语言描述,欧几里得的目的是解方程 $x^2 = c$.首先他应用卷 1 命题 45 作出一个面积为 c 的矩形(图 2.27),再把矩形的边 BE 和 EF 放在一条线段上,然后作 BF 的二等分点 G,再作以 GF 为半径的半圆 BHF,其中 H 是过 BF 上的点 E 的垂线与半圆的交点.由卷 2 命题 15 可知,以 BE 和 EF 为边的矩形与 EG 上的正方形的和等于 GF 上的正方形,又因为 $GF = GH$,GH 上的正方形等于 GE 和 EH 上的正方形的和,所以 EH 上的正方形满足问题的条件,即为要作的正方形.虽然我们可以把命题 14 看作是解二次方程,但把这一证明和卷 2 命题 5 的应用看作是代数似乎有些牵强.

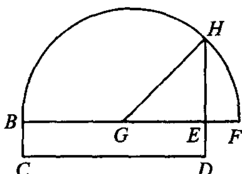


图 2.27　《原本》卷 2 命题 14.

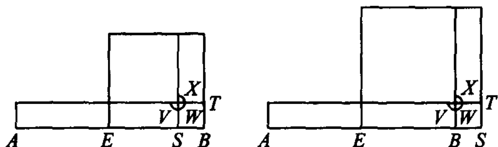


图 2.28　《原本》卷 6 命题 28、29.

最后,我们给出卷 6 中的一些与这一争论相关的例证.

命题 Ⅵ－28　在一条已知线段上作一个等于已知直线形的平行四边形,它是由去掉一个与已知平行四边形相似的图形后形成的;这个已知的直线形必须不大于在原线段一半上的平行四边形且这个平行四边形与去掉的平行四边形相似.

命题 Ⅵ－29　在一条已知线段上作一个等于已知直线形的平行四边形,且在这条线段的延长线上有一个平行四边形与已知的平行四边形相似.

在卷 1 的命题 44 中已经出现了在已知直线上作平行四边形的概念了,在上述问题中,欧几里得讨论了"去掉(deficient)"和"加上(exceeding)"的应用.在第一种情况,作一个面积已知、一底边小于已知线段 AB 的平行四边形.在去掉的线段 SB 上的平行四边形与已知的平行四边形相似(图 2.28);在第二种情况,他作一个面积已知、底边大于已知线段 AB 的平行四边形,在加上的线段 BS 上的平行四边形与已知的平行四边形相似.在第三章圆锥曲线的讨论中将明显地体现出这种"去掉"和"加上"思想的重要性.为了讨论几何代数,最简单的方法是假设已知的平行四边形为正方形,我们将在这种特殊情形下把他的命题和证明表示成代数的形式,这样两种情况下所作的平行四边形一定是矩形.

如果用 b 表示 AB,已知直线形的面积为 c,问题可化简为在 AB(命题 28)或 AB 的延长线(命题 29)上找一点 S,使得 $x = SB$ 分别满足条件:第一种情况,$x(b - x) = c$;第二种情况,$x(b + x) = c$,也就是说,需要分别解两个二次方程 $bx - x^2 = c$ 或 $bx + x^2 = c$.欧几里得在两种情况都先确定 AB 的中点 E,然后在 BE 上作面积为 $(b/2)^2$ 的正方形,对第一种情况,选择点 S 使得 ES 上的正方形面积为 $(b/2)^2 - c$,这便是命题中为什么要提出那样的条件,实际上即是 c 不能大于 $(b/2)^2$,由点 E 的选择可以得出

$$x = BS = BE - ES = \frac{b}{2} - \sqrt{\left(\frac{b}{2}\right)^2 - c}.$$

对第二种情况,选择的点 S 使得以 ES 为边的正方形的面积是 $(b/2)^2 + c$,那么有

$$x = BS = BE - ES = \sqrt{\left(\frac{b}{2}\right)^2 + c} - \frac{b}{2}.$$

对这两种情况,欧几里得证明了要作矩形等于曲尺形 XWV 同时 XWV 又等于已知的面积 c,从而说明了他对点 S 的选择是正确的.若用代数的形式表示,第一种情况实际上就是

$$x(b-x)=\left(\frac{b}{2}\right)^2-\left[\left(\frac{b}{2}\right)^2-c\right]=c.$$

第二种情况就是

$$x(b+x)=\left[\left(\frac{b}{2}\right)^2-c\right]-\left(\frac{b}{2}\right)^2=c.$$

几何代数实际上是源于对巴比伦结果的几何翻译,特别是在一些特例中考虑到这种几何过程与代数过程的相似性时,这一观点更是令人信服.因此我们可以说,希腊人将问题翻译成几何形式,并给出必要的证明,这可能与他们发现并不是每一条线段都可以用"数"来表示有关.我们还可以进一步说,一旦把问题译为几何的形式,人们就可以对平行四边形也像矩形那样来表述和证明相关的结果,二者并无本质的差异.支持传播和交流看法的进一步的论点是巴比伦最初的方法本身可能已经含有"朴素"的几何形式,而这种形式非常适合转换成更复杂的希腊几何.正如我们前面提到,欧几里得的几个命题看上去与巴比伦解方程的算法的形式是等价的.

希腊数学家与巴比伦的数学书记员有直接联系吗?长期以来,有些学者们认为这种联系的可能性不存在,因为从公元前6到公元前4世纪,并没有任何关于巴比伦数学的记载,而这种联系只可能在这段时期内发生,而且当时希腊的包括数学家在内的上层(优秀)人物十分鄙视那些书记员的工作,因为这些人在古巴比伦时期并不属于社会精英.但最近的发现表明,在公元前1纪中期,那里确实还有数学活动,而且那时的美索不达米亚语言使用的是一套新的字母,并且变成用墨水写在纸草书上,而泥板书上书写的楔形文字便成了值得保存的重要文献,于是可以从事这一工作的人员就上升为精英,变成了对行使国家职能具有重要作用的、身怀绝技的专家.另外,从公元前6世纪开始,美索不达米亚是与希腊保持联系的波斯帝国的一个城邦.

尽管有许多学者争论着巴比伦数学可能已经被译成了希腊几何,但仍没有直接的证据来说明在公元前4世纪或之前巴比伦数学已传播到了希腊,特别是没有证据说明希腊人曾使用过任何形式的数值代数.甚至毕达哥拉斯学派的数论也是以某种几何形式来表述的.也许可以这样说,即使希腊人应用了我们认为的代数过程,而他们的那种几何化的数学思想使得所有这些过程都自然地表示成了几何形式.公元前300年之前的希腊人并没有代数记号,因此他们对量的表达式进行运算的惟一方法是把它们表示成几何的形式.实际上,希腊数学家们非常擅长几何操作.最后我们指出,除了几何方法外,希腊人没有其它办法来表示二次方程的无理数解.

巴比伦的代数在公元前4世纪是否以某种形式传播到希腊?我们本节中讨论的定理是否可以看作是"代数"?这里仍不能给出此类问题明确的答案.对此有兴趣的读者可以查阅后面列出的文献或认真地阅读原始文献.

2.2.4 圆和五边形作图

在卷1和卷2中,欧几里得研究了直线形——由线段围成的图形的性质,卷3中,他开始讨论最基本的曲线——圆的性质.希腊人对圆情有独钟,不论怎样旋转,它总是保持一样,他们认为圆是最完美的图形.同样,他们也考虑三维的圆,即球,认为球是最完美的立体图形.这些哲学思想成为希腊天文学思想的基础,我们将在第4章对此予以讨论.

《原本》卷3和卷4中的许多定理可以追溯到希腊数学的早期,例如希波克拉底关于月牙形的研究中已经清楚地认识到圆的重要性.卷3中的各个命题相互独立,但从内容上看,本卷是按与圆

相关的正多边形作图来组织安排的,即作内接或外切于圆的正多边形,这些内容在卷 4 中得到进一步完善. 特别地,卷 3 前半部分中的大多数命题被应用于卷 4 中最复杂的作图问题 —— 作正五边形. 三角形、正方形和六边形的作图相对要直观些,而这些较直观的图形的作法可能毕达哥拉斯学派已经给出过. 但五边形的作图中却包括一些新的概念,如把一线段按"中外比"分为两部分,因此这很可能是后来的发展,多半是公元前 4 世纪前期蒂奥泰斯的工作. 这一作图方法本身在《原本》卷 8 中关于多面体的一些作图问题中又获得应用.

补遗 2.2	卷 3 中的定义
	2. 与圆相切的直线是和圆接触,但延长后并不与圆相交的直线. 3. 弓形是由一条直线和一段圆弧所围成的图形. 8. 在弓形弧上取一点,该点与弓形底上的两个端点连线所成的角是弓形角.

在给出一些相关定义后(补遗 2.2),欧几里得在卷 3 中开始讨论一些基本的作图和命题,然后他接着给出圆的切线的作法.

命题 Ⅲ - 16 过圆直径的端点作与直径垂直的直线,则该线落在圆外,且在圆周和此直线间再不能插入其它的直线.

命题中说过直径的端点与直径垂直的直线即现在所称的**切线**,欧几里得只在一个推论中作了如定义 2 所说的"接触"圆的注释,但是,在圆周和直线间不能再插入其它的直线这一说法在微积分引入之前就成了切线定义的一部分. 欧几里得用反证法证明了这一命题.

命题 18 和 19 给出了命题 16 的部分逆命题,前者是说,过圆心和切点的直线与切线垂直,后者是说过切点与切线垂直的直线一定过圆心. 命题 20 和 21 也给出了类似的结论,同弧上的圆心角是圆周角的二倍,同一弓形上的角彼此相等,从图 2.29 可以明显得出这两个命题和命题 22 的证明,命题 22 是说内接于圆的四边形其对角的和等于两直角.

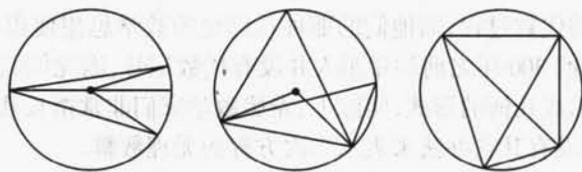


图 2.29 《原本》卷 3 命题 20、21、22.

命题 31 指出,半圆所对的角是直角. 如果把平角看作是一个角,那么由命题 20,我们可以马上得出命题 31 的结果,既然半圆对的圆周角是直径所成的平角的一半,反之直径所成的平角等于两个直角. 但是欧几里得并没有把平角看作是一个角,因此他给出了一个不同的证明.

我们下面介绍命题 36 和 37,为五边形的作图做准备.

命题 Ⅲ - 36 过圆外一定点分别作圆的一条切线和一条割线,那么由整个割线段与其落在圆外的线段所围成的矩形等于切线上的正方形.

这一命题也许会令读者想到卷 2 命题 6,实际上此命题的证明中应用了卷 2 命题 6. 这里我们只考虑较简单的情况,设割线 $DCFA$ 过圆心 F(图 2.30),连接 FB 后得一直角 FBD,由卷 2 命题 6 可知,

AD 和 CD 构成的矩形与 FC 上的正方形的和等于 FD 上的正方形，但 $FC = FB$，且 FB 和 BD 上的正方形的和等于 FD 上的正方形，因此 AD 和 CD 构成的矩形等于 DB 上的正方形，命题得证. 割线不过圆心的情况要稍复杂一些.

命题 37 是 36 的逆命题，即过圆外一点向圆引两条直线，其中一条是圆的割线，另一条落在圆上，如果命题 36 中的那种矩形和正方形之间的关系成立，那么落在圆上的直线与圆相切.

图 2.30 《原本》卷 3 命题 36.

欧几里得在讨论了圆内接三角形、正方形，三角形和正方形的内切圆，圆外切三角形、正方形，三角形、正方形的外接圆等简单技巧后，他在卷 4 中开始研究五边形的作图问题. 欧几里得正五边形的作图过程分为两步，第一步是作底角等于顶角二倍的等腰三角形(卷 4 命题 10)，第二步作圆的内接五边形(卷 4 命题 11). 与通常一样，欧几里得并没有说明他是怎样发现这个作法的，但认真研读后我们可以得到他分析问题的一条线索，因此我们假定图形已经作出，并且尽力发现他的思路.

设 $ABCDE$ 是圆的内接正五边形(图 2.31)，作对角线 AC 和 CE，因为 $\angle CEA$ 和 $\angle CAE$ 所对的圆弧分别是 $\angle ACE$ 所对圆弧的二倍，所以 ΔACE 是底角等于顶角二倍的等腰三角形，因此可以把五边形的作图归结为三角形的作图. 在 ΔACE 中，作 AF 二等分 $\angle A$，而 ΔAFE 与 ΔCEA 相似，所以 $EF : AF = EA : CE$，但 ΔAFE 和 ΔAFC 都是等腰三角形，所以 $EA = AF = FC$，因此 $EF : FC = FC : CE$，或用现代的表示是 $FC^2 = EF \cdot CE$. 那么作图问题就归结为在已知线段 CE 上找一点，使得 CF 上的正方形等于 EF 和 CE 构成的矩形，而这正是卷 2 命题 11 的作图，一旦找到了点 F，那么，分别以 C 为圆心、CE 为半径和 E 为圆心、CF 为半径画两个圆，它们的交点 A 就是要作的三角形的第三个顶点，即可作出底角等于顶角二倍的等腰三角形.

图 2.31 正五边形的作图.

图 2.32 《原本》卷 4 命题 10.

欧几里得在卷 4 命题 10 给出了这一作图过程(图 2.32)，但是，正如毕达哥拉斯定理的证明一样，他在证明中没有应用相似性的方法，而是用了其它的方法. 目的是证明 $\alpha = 2\beta$，如果可证明 $\beta = \delta$，那么 $\beta + \gamma = \delta + \gamma = \epsilon$，又因为 $\alpha = \beta + \gamma$，所以 $\epsilon = \alpha$，但 $AE = AF$，通过作 $AE = FC$ 即可得 ΔAFC 是等腰三角形，即 $\delta = \gamma$. 最后，即可得出 $\alpha = \beta + \gamma = \delta + \delta = 2\delta$. 要证明 $\beta = \delta$，可作 ΔAFC 的一个外接圆，因为 CE 和 EF 所构成的矩形等于 FC 上的正方形，所以这个矩形也等于 AE 上的正方形，由卷 3 命题 37 可得出 AE 是圆的切线，而卷 3 命题 32 指出，如果一条直线与圆相线，

且由切点作一条过圆内部的直线与圆相截，那么直线和切线所夹的角等于截线所对的另一弓形上的角，这里切线是 AF，截线是 AE，所以可证 $\beta = \delta$. 这就完成了整个作图过程的证明.

作出了等腰三角形后，那么再作圆内接正五边形就很容易了. 欧几里得首先作一圆内接等腰三角形 ACE，然后作三角形中 $\angle A$ 和 $\angle E$ 的平分线，它们与圆周分别交于 D 和 B 点，A、B、C、D、E 即是正五边形的五个顶点.

欧几里得在卷4的最后作了圆内接正六边形和正15边形，但并没有提及其它正多边形的作图，可能他已经知道了从已经作出的多边形开始，作圆内接 $2^n k$($k = 3,4,5$) 边形是容易的；通过与正15边形作图类比，他甚至可能已经知道：如果可以作出圆内接 k 边形和 l 边形(k 与 l 互素)，那么就能够容易地作出 kl 边形. 然而，七边形的作图最早出现在阿基米德的著作中，我们并不清楚欧几里得是否知道七边形的作图，但对欧几里得来说，因为这种作图除了圆规和直尺外还需要其它工具，所以这是高等数学的内容，而非"初等"的.

2.4.5 比率和比

《原本》前四卷中一些重要的结论本可以通过相似的基本思想得到证明，但欧几里得在组织其内容时尽可能地不用相似概念，这也可能是因为相似的思想首先需要用到量的相同比这一很微妙的概念，实际上，卷5集中研究了比例的基本概念，而这些概念是卷6中研究相似性的前提.

卷5对数学的近代发展方向起了重大的影响，许多世纪以来，欧几里得的比例理论决定了如方程理论、分数理论、实数系统等诸多研究的实质. 因此，理解欧几里得的原著对理解其后数学的发展历史也是非常重要的.

卷5中的前两个定义比较容易理解(见补遗2.3). 这里值得注意的是欧几里得研究了量. 卷5是研究量(即连续量)的比例理论，卷7研究数(即离散量)的比例理论. 尽管在现代数学里，可以容易地把后者归入到前者中，但当时在欧几里得看来，这是不可能的，因为这两个概念互不相同，所以必须分别对它们进行研究. 无论如何，在卷5的前两个定义中，用一个量去度量另一个量的意思是指第一个量的(正)整数倍等于第二个量.

补遗 2.3	卷5中的定义
	1. 如果一个较小量可以度量一个较大量时，我们把较小量叫做较大量的**部分**. 2. 如果一个较大量可以被一个较小量度量时，我们把较大量叫做较小量的**倍量**. 3. 两个同类量之间的一种数量关系叫做**比**. 4. 如果一个量增大几倍后可以大于另一个量，则说两个量**有一个比**. 5. 有四个量，第一量比第二量与第三量比第四量叫做有**相同比**，如果对第一和第三量取任何相同倍数，又对第二和第四量取任何相同倍数，而当第一与第二倍量之间依次有大于、等于或小于的关系时，第三与第四倍量之间也有相应的关系. 6. 有相同比的四个量叫做**成比例的量**. 7. 在四个量之间，第一、第三两个量取相同倍数，第二、第四两个量取另一相同倍数，若第一个倍量大于第二个倍量，且第三个倍量不大于第四个倍量时，则说第一个量与第二个量的**比大于**第三个量与第四个量的比. 9. 当三个量成比例时，则说第一量与第三量的比是第一量与第二量的**二次比**. 10. 当四个量成连比例时，第一量与第四量的比叫做第一量与第二量的**三次比**.

虽然第三和第四个定义有些模糊，但它们有助于我们理解欧几里得的比的含义. 两个量之间有一个比，仅当这两个量是同类的，如两条线、两个面或两个立体，且其中的一个与另一个之间存在倍数关系. 例如，因为圆周和切线间的角的倍数总不能超过已知直线角，所以这两个角之间不存在比.

卷5的中心概念是定义5中的**相同比**. 现在，一般把等式 $a:b=c:d$ 看作是与两个分数 a/b 和 c/d 相等等价，但是欧几里得时期或更早期的希腊正式著作中根本没有使用分数，这使我们想到前面说的单位不能再分. 在计算著作中，分数自然是存在的，但希腊人更常用的是传统的埃及单位分数，且称之为"部分". 现存有许多类似于兰德纸草书中的"除法" 表，说明了怎样用这些"部分" 来计算. 例如，有一个表中写着："对12，第17部分是 $\frac{1}{2},\frac{1}{12},\frac{1}{17},\frac{1}{34},\frac{1}{51},\frac{1}{68}$，" 这在今天的计算中是以 $\frac{12}{17}$ 表示.[19] 一般地，在希腊正式的数学中，普通分数是用两个量的比来替代. 而定义两对数的相同比则比较容易，欧几里得在卷7的定义20(见补遗2.5) 中给出了这种定义. 然而回顾在公元前5世纪后期，人们已发现没有办法把正方形的边和对角线都表示成数，于是希腊数学家便寻求一种可以应用到所有的量(包括不可公度量) 的定义. 目前可以肯定的是，寻求这种定义的过程是从欧几里得算法即我们熟知的求两数的最大公因数的算法开始的.

卷7命题1和2给出了在欧几里得之前就已为人们所知道的这一算法. 已知两个数 a 和 b，且 $a>b$，若从 b 中连续减去 a 后，得一余数 c(当然 $c<b$)，然后再从 b 中连续减去 c，如果这样一直做下去，最终要么是得到一个可以"量尽" 它前面数的数 m(卷7命题2)，要么是得到一个单位(1)(卷7命题1). 对第一种情况，欧几里得证明了 m 是 a 和 b 的最大公度量(即最大公因数). 对第二种情况，他证明了 a 和 b 互素. 例如，已知两个数18和80，按照上面的算法做4次后得余数8，再从18中连续两次减去8，得余数2，最后，从8中连续4次减去2，余数为0，所以2是18和80的最大公因子. 这种算法也说明可以把80和18的比表示成(4,2,4) 的形式，这意味着若将这一算法应用到其它满足 $a:b=80:18$ 的数时 a，b 时也将得出(4,2,4).

> **人物小传**
>
> **蒂奥泰德(Theaetetus)(公元前417—369)**
>
> 柏拉图曾有一段与蒂奥泰德的对话，从中我们可知他的一些生平. 他出生于雅典附近的一个富裕家庭并在那里接受教育. 他20岁前与狄奥多罗斯(Theodorus of Cyrene) 的一次相会使他走上了数学研究的道路. 狄奥多罗斯不仅向他讲了 $\sqrt{2}$ 与1是不可公度的，还给他讲了不超过17的非平方数的平方根. 然后蒂奥泰德开始研究不可公度问题(在赫拉克利亚(Heraclea)，公元前375年后在雅典). 公元前369年他应征入伍，在科林斯(Corinth) 的一次战争中受伤，之后不久因痢疾而死.

很可能是蒂奥泰德首先研究了将欧几里得算法应用于这样的数量的可能性，从而开辟了量的比例定义发展的新阶段. 其结果作为卷10命题2和3而出现，即通常称之为"互易减法"(anthyphairesis) 的过程. 蒂奥泰德证明了怎样判断两个量 A 和 B 是否有一个公共度量(可公度) 或没有公共度量(不可公度)，其过程与对数的情形(最大公因数的判定) 基本相同. 假定 $A>B$，先从 A 中连续减去 B，假设减 n_0 次后得到一个小于 B 的余量 b，再从 B 中连续减去 b，假设减 n_1 次后得到一个小于 b 的余量 b_1，欧几里得在卷10命题2中指出，如果这个过程永远不能结束，那么最初的两个量是不可公度的；另一方面，如果最后得出的量可以量尽它的前一个量，那么这个量就是最初两个量的最大公度量(卷10命题3). 这里自然产生一个问题，怎样判断这一过程能否结束呢？一般来说，这种判断是很困难的，但在一些特例中，通过对余量重复方式的考察，可以证明过程的没有

终结.

例如,证明正方形的边 s 和对角线 d 是不可公度的,首先从 $d+s$ 中两次减去 s,余数 $s_1=d-s$(图2.33).因为 $s=s_1+d_1$,那么接下来从 s_1+d_1 两次减去小正方形的边 s_1,余数 $s_2=d_1-s_1$,也就是说,在每一步中我们都可以从边和对角线的和中两次减去正方形的边.因为这一过程永远不能结束,所以 $d+s$ 和 s 是不可公度的,从而 d 和 s 也是不可公度的.一些历史学家认为这一过程最早发现了 d 和 s 的不可公度性.

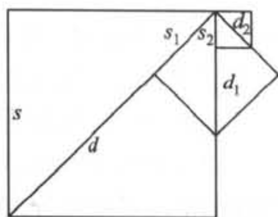


图2.33 正方形边和对角线的不可公度性(第二种可能).

已知互易相减过程,蒂奥泰德能够给出对所有的量都适用的相同比的定义.设有两对量 A、B 和 C、D,对每对量应用互易相减法可得到两个等式序列:

$$
\begin{array}{llll}
A=n_0B+b & (b<B) & C=m_0D+d & (d<D)\\
B=n_1b+b_1 & (b_1<b) & D=m_1d+d_1 & (d_1<d)\\
b=n_2b+b_2 & (b_2<b_1) & d=m_2d+d_2 & (d_2<d_1)\\
 & \cdots & & \cdots\\
 & \cdots & & \cdots
\end{array}
$$

如果这两个数列$(n_0,n_1,n_2,\cdots)$、$(m_0,m_1,m_2,\cdots)$逐项都相等,且同时结束于(例如)$n_k=m_k$,那么我们可以算出 $A:B$ 和 $C:D$ 都等于某个相同的整数比.这就使蒂奥泰德得以给出一般定义,即:如果(可能永不结束的)数列$(n_0,n_1,n_2,\cdots)$和$(m_0,m_1,m_2,\cdots)$逐项相等,那么 $A:B=C:D$.当然对一些特殊的量来说,这个定义颇不好用,但是在一些有趣的例子中,确定 $n_0,n_1,n_2,\cdots$ 要相对容易些.例如,我们已经讲到了正方形的对角线和边的比可以用$(1,2,2,2,\cdots)$表示.虽然亚里士多德曾说,蒂奥泰德的相同比的定义是当时被大量应用的一个定义,但《原本》卷5中出现的却是欧多克斯给出的另一个定义[20].

人物小传	欧多克斯(408 B.C.—355 B.C.)
	欧多克斯年轻时在小亚细亚海岸附近的尼多斯(Cnidus)学习医学.当他访问雅典时,被雅典学派关于哲学和数学的演讲吸引,之后开始研究哲学和数学.后来,他又访问了埃及并在那里进行了天文观测和埃及历法的探讨.当他返回家乡后,开办了一所学校并从事自己的研究,虽然他曾和他的学生一起返回雅典,但他后来的大部分时间都是在尼多斯度过.他的几何研究和把球面几何应用于天文学的杰出工作使他闻名于后世.

虽然我们不知道是什么激发欧多克斯给出相同比的一个新定义,但可以作一个合理的推测.比如,欧多克斯的定义说,如果 $A:B=C:D$,那么 $A>n_0B$,而 $C>m_0D$(因为 $n_0=m_0$).又因为 $n_1A=n_1n_0B+n_1b=(n_1n_0+1)B-b_1$,即 $n_1A<(n_1n_0+1)B$,同样 $n_1C<(n_1n_0+1)D$,把 A 和 B 的倍数与 C 和 D 的相应的倍数作比较后可得,对不同的数 r、s,只要 $rC>sD$,总有 $rA>sB$ 和只要 $rC<sD$,总有 $rA<sB$.这样,欧多克斯给出了相同比和比例的定义(即卷5定义5和定义6)并用它们严格地证明了卷5中的许多比例定理,从而为卷6中的相似性的几何理论奠定了坚实的基础.

如果用代数符号表示,定义5是说:$a:b=c:d$,如果对于任给正整数 m、n,只要 $ma>nb$,总有 $mc>nd$;只要 $ma=nb$,总有 $mc=nd$;只要 $ma<nb$,总有 $mc<nd$.这一定义的另一种现代表示是,对任一个分数 n/m,商 a/b 和 c/d 同时大于、等于或小于这个分数.

欧几里得给出的定义 9 现今叫做比的平方,或等价于平方比.他指出,如果 $a:b=c:d$,那么 $a:c$ 是 $a:b$ 的二次比,表示成现代的形式即是 $a:c=(a:b)(b:c)=(a:b)(a:b)=(a:b)^2=a^2:b^2$,或分数形式 $a/c=(a/b)^2=a^2/b^2$.但是,正如没有把量相乘一样,欧几里得没有把比相乘也没有把分数相乘,他只是用数来乘以一个量,同样他也没有把量相除.因此我们不能把欧几里得的比 $a:b$ 译成一个分数的形式来进行标准的算术运算.另一方面,欧几里得把两个量的二次比和它们的平方比或三次比与它们的立方比看作是等价的,这样就可以清楚地说明量的"平方"的意义.

卷 5 的第一个命题用现代的符号描述为:如果 $ma_1,ma_2,\cdots,ma_n$ 是 $a_1,a_2,\cdots,a_n$ 的同倍量,那么 $ma_1+ma_2+\cdots+ma_n=m(a_1+a_2+\cdots+a_n)$.类似地,卷 5 命题 2 实际是说 $ma+na=(m+n)a$,命题 3 即是 $m(na)=(mn)a$.也就是说,卷 5 的前几个命题给出了现代的分配律和结合律.

引出相同比的第一个命题是卷 5 命题 4,它是说如果 $a:b=c:d$,那么 $ma:nb=mc:nd$,这里 m,n 是任意的数.要证这一等式,欧几里得必须先证明:若 $p(ma)$、$p(mc)$ 是 ma 和 mc 的同倍量,$q(nb)$、$q(nd)$ 是 nb、nd 的同倍量,那么根据 $p(ma)>=<q(nb)$,同样可以判定 $p(mc)>=<q(nd)$.因为 $a:b=c:d$,以及根据分配律和最初量的相同比的定义,欧几里得得出了倍量的同比性.

命题 5 和 6 是把命题 1 和 2 中的加号换成减号.命题 7 证明了如果 $a=b$,那么 $a:c=b:c$,$c:a=c:b$.命题 8 是说如果 $a>b$,那么 $a:c>b:c$,$c:b>c:a$,欧几里得应用定义 4 和 7 证明了命题 8 的前一部分结论.因为 $a>b$,所以存在正整数 m,使得 $m(a-b)>c$(定义 4).设 q 是使 c 等于或大于 mb 的第一个倍数,那么 $qc\geqslant mb>(q-1)c$.因为 $m(a-b)=ma-mb>c$,所以 $ma>mb+c>qc$.又因为 $mb<qc$,由定义 7 可知 $a:c>b:c$.类似地可以证明第二个结论.

卷 5 命题 11 给出了比例的传递律(transitive law),即如果 $a:b=c:d$ 且 $c:d=e:f$,那么 $a:b=e:f$,命题 16 是说如果 $a:b=c:d$,那么 $a:c=b:d$.其它一些命题给出了成比例的量的其它一些性质,特别是涉及了在各种比例式中给比例前项或后项加上或减去一个数量的情况.

2.4.6 相 似

卷 6 中应用量的比例理论中的结论证明了相似直线形的理论,本卷给出了相似直线形的第一个定义(见补遗2.4).因为卷6中引用了公元前5世纪的参考文献,所以一般认为毕达哥拉斯学派已经知道了卷6中的结论.但是,相似的思想基础——等比的概念最初是建立在所有的量都可以被看作是数的思想基础上的.一旦这一概念不正确,那么这些结论也就失去了存在的基础.这并不是说数学家们不再使用这些结论.即使不能给出一个形式的定义,他们直觉上还是认同等比的概念.与近代一样,希腊时期的数学家经常忽视基础问题而去追求新结论的发现.明智的数学家知道,数学的基础最终会得到加强.约公元前 360 年,当这一情况一旦来临,那些相似性的结论终于能被组织成合乎逻辑的、可接受的体系.我们并不知道谁最后作了这一组织工作,但也许可以肯定的是,除了本卷命题 1 的证明外,其它内容均很少能作改动,而命题 1 是惟一直接以欧多克斯的定义为基础的一个命题.

	卷 6 定义
补遗 2.4	1.若直线形的角对应相等且夹等角的边对应成比例,那么称它们是**相似直线形**. 3.分一线段为二线段,当整体线段比长线段等于长线段比短线段时,则称此线段被分成**中外比**. 4.一个图形中,由顶点到底边的垂线叫做图形的**高**.

命题 Ⅵ - 1 等高的三角形或平行四边形的比等于它们底的比.

已知 ΔABC 和 ΔACD 等高,欧几里得要证明 BC 和 CD 的比等于 ΔABC 和 ΔACD 的比.按欧多克斯定义的要求,欧几里得把 BD 两端延长,使得能够在这条延长线上作出 BC 和 CD 的任意倍数(图 2.34).虽然他提到了"任意条"线段,但他没有这样的表示记号,因此只是取了两条线段.大概欧几里得认为这是一种我们所谓的"可推广举例",他在其它命题的证明中也使用了类似的手法.因此,他从每一边的两条线段入手,因为等高等底的三角形相等,所以若底 HC 是 BC 的几倍,那么 ΔAHC 也是 ΔABC 的几倍.这对 ΔALC 和 ΔACD 也同样成立.又因为 ΔAHC 和 ΔALC 同高,所以当 HC 大于、等于或小于 CL 时,ΔAHC 也大于、等于或小于 ΔALC.若分别取 ΔABC 及其底 BC 的同倍量,ΔACD 及其底 CD 的同倍量,那么按欧多克斯的定义比较这两结果就可得要证的结论:$BC : CD = \Delta ABC : \Delta ACD$.因为平行四边形是一个三角形的两倍,所以同样的结论对平行四边形也是成立的.

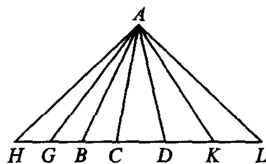


图 2.34 《原本》卷 6 命题 1.

命题 2 证明了平行于三角形一边的直线截得三角形另两边的线段成比例.命题 3 是说三角形一角的角分线分对边所成的两条线段的比等于三角形另两边的比.接着,欧几里得给出两个三角形相似的不同条件.因为相似的定义要求两三角形的对应角相等且对应边成比例,所以欧几里得证明了这两个条件中只要有一个成立,就可判定两个三角形相似.他还给出了只有一对角相等且有两对边成比例时两个三角形相似的条件.命题 8 证明了过直角三角形的直角顶点作斜边的垂线后得到的两个小直角三角形与原三角形相似.

卷 6 中包括一些求比例项的作图.已知线段 a、b、c,欧几里得给出了确定分别满足 $a : b = b : x$(命题 11)、$a : b = c : x$(命题 12) 和 $a : x = x : b$(命题 13) 的 x 的方法,其中最后一个结论等价于求平方根 —— 即求解 $x^2 = ab$ —— 因此这一结论也等同于卷 2 命题 14.实际上,它们的证明结构相同,只是前者应用相似而后者应用几何代数来证明问题,

命题 16 本质上与下述命题相似:比例中项的积等于两比例外项的积.但是因为欧几里得从来没有把量相乘,他不能按卷 5 的内容表述这一结论.但在卷 6 的几何问题中他给出了等价于乘法的运算,但只是对线段而言.

命题 Ⅵ - 16 如果四条线段成比例,则两外项构成的矩形等于两内项构成的矩形;如果两内项构成的矩形等于两外项构成的矩形,那么这四条线段成比例.

命题 19 在后来显得非常重要,它解释了欧几里得二次比的概念.

命题 Ⅵ - 19 相似三角形的比等于其对应边的二次比.

如果用现代的语言描述,那么可以把这一结论中的"二次比"换成"比的平方".欧几里得没有把量或比相乘,因为比不是数量,他也没有把比看作是数.因此,对这一特殊的命题,欧几里得需要在 BC 上作一点 G,使得 $BC : EF = EF : BG$(图 2.35),那么 $BC : BG$ 是 $BC : EF$ 的二次比(图 2.33).为了证明这一结论,他证明了 ΔABG 和 ΔDEF 相等.又因为 ΔABC 和 ΔABG 的比等同于 BC 和 BG 的比,即可得出要证的结论.命题 20 把这一结论推广到相似多边形上,特别地,两条线段的二次比就等于线段上正方形的比.

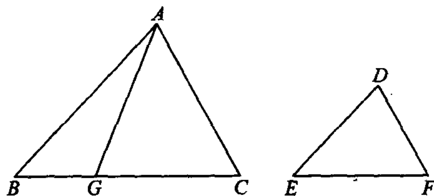


图 2.35 《原本》卷 6 命题 19.

当然,不相似的两个平行四边形的对应角也可以是相等的,欧几里得对此类问题也进行了研究,但只是使用了一个未正式定义的概念:

命题 Ⅵ - 23 对应角相等的平行四边形的比等同于边的比的复合比(the ratio compounded of the ratios).

由该命题的证明可以理解欧几里得所谓"复合"的意思,至少在线段比的情形是如此.已知两个比 $a : b$ 和 $c : d$,先作一线段 e,使得 $c : d = b : e$,那么 $a : b$ 和 $c : d$ 的**复合比**即是 $a : e$.用现代符号表示,分数 a/e 只是分数 a/b 和 $c/d = b/e$ 的简单相乘.有趣的是,虽然欧几里得除了此处以外再没有考虑复合比,但这一概念在希腊后期和中世纪却变得非常重要.

2.4.7 数 论

《原本》卷 7 是三卷研究初等数论内容中的第一卷,卷 7、8 和 9 构成了一个完全独立的体系,这三卷中并没有提到前六卷中的任何内容,只是在卷 9 中才有一些和前面几何内容相联系的算术内容.欧几里得在卷 7 开始给出了一个例证,说明他保留了亚里士多德对量和数的清晰的区分.前 6 卷主要讨论了量,特别是长度和面积,卷 5 主要讨论了量的一般理论.但欧几里得在卷 7—9 中只讨论了数,他没有把它们看作是某种类型的量,而是完全独立的实体.因此,虽然卷 7 中有许多结果只是卷 5 中结论的特例,但欧几里得认为他们是完全不同的.欧几里得在这几卷中用线段表示数可能会对我们产生误导,但他在具体的证明中却没有使用这种表示,可能这种表示法对他来说是绝无仅有了.

算术卷中的许多命题可在毕达哥拉斯学派时期的算术著作中找到,但从《原本》卷 10 中对卷 7 内容的使用来看,其中复杂的细节好像都来自同一数学家,即蒂奥泰德的工作.蒂奥泰德研究了毕达哥拉斯学派结构松散的数的理论后,通过引入精确的定义和详细的证明来使数的理论严密化,而这些定义和证明被欧几里得纳入到自己的著作中.

补遗 2.5	卷 7 的定义
	1. 每一个事物是作为一个**单位**而存在,并称之为 1.
	2. 一个**数**是由多个单位合成的.
	3. 一个较小数是较大数的**一部分**,如果它可以量尽较大数.
	4. 一个较小数是较大数的**几部分**,如果它不能量尽较大数.
	5. 较大数若能被较小数量尽,则它是较小数的**倍数**.
	11. **素数**是只能被单位量尽的数.
	12. **互素的数**是只能被作为公度的一个单位所量尽的几个数.
	15. 所谓一个数**乘**一个数,就是被乘数自身相加多少次得出的数,这个次数是另一个数中单位的个数.
	20. 当一个数是第二个数的某倍、某一部分或某几部分,第三数是第四数的同倍、同一部分或同几部分,称这四个数是**成比例的**.

卷 7 与其它卷一样,也是从定义开始(补遗 2.5).与卷 1 的第一个定义一样,卷 7 第一个定义从现代观点讲是没有意义的,但在欧几里得看来,该定义是作为"事物"这一概念的数学抽象出现的.更有趣的是第二个定义,一个数是多个单位,因为"**多个**(multitude)"意思是复数,而单位不是复数,在欧几里得和早期的毕达哥拉斯学派看来,1 不是一个数.

定义 3 和 5 实际上是卷 5 定义 1 和 2 的重复，定义 4 对任意量来说是毫无意义的. 如果注意到欧几里得认为一个数不能度量其自身，那么定义 11 和 12 基本上是现代素数和互素的定义. 定义 15 有些奇怪，它是欧几里得惟一一个有关算术运算的定义，其中他假定加法和减法是已知的. 除此之外，再没有与卷 5 中相类似的定义.

卷 7 的前两个命题主要讨论了欧几里得算法. 随后的几个命题与卷 5 中的相似，如欧几里得在命题 5 和 6 中相当于证明了分配律 $\frac{m}{n}(b+d)=\frac{m}{n}b+\frac{m}{n}d$. 他在卷 5 命题 1 中对量已经证明了这一命题，只不过那里讨论的是(整数) 倍而不是卷 7 中的部分(这里表示分数). 这些证明实际上是等价的. 欧几里得不是简单地引述卷 5 中的结论，这也说明他并不把数看作是一种量.

命题 11 – 22 包括许多关于数的比例的标准结果，而他在卷 5 中讨论量比例时已经证明了其中一些结果. 这些命题大部分又被用在随后的两卷中. 特别地，命题 16 证明了乘法交换律. 命题 19 给出了比例的一般验证法，即当且仅当 $ad=bc$ 时，$a:b=c:d$. 欧几里得对线段已经作过类似的证明(卷 6 命题 16)，但是对数来说，这个证明是非常困难的. 已知 $a:b=c:d$，可得 $ac:ad=c:d=a:b$，同样，$a:b=ac:bc$，于是 $ac:ad=ac:bc$，所以 $ad=bc$. 类似地可以证明它的逆命题. 命题 20 是说，如果 a,b 是比 $a:b$ 中最小的数，那么 a 和 b 可以各除 c 和 d 同样的次数，这里 $c:d=a:b$. 由此可得互素的数是有相同比的数中最小的数，反之也成立.

命题 23 – 32 进一步讨论了素数和数的互素. 特别地，它们给出了欧几里得的整除性理论，同时与卷 9 命题 14 一起给出了算术基本定理，即任何数都可表示成惟一的素数积的形式.

命题 Ⅶ – 31　任一合数可被某个素数量尽.

命题 Ⅶ – 32　任一数或者是素数或者可被某个素数量尽.

后一个命题是前一命题显然的推论. 命题 31 本身则可以用欧几里得在算术卷中常用的技巧——最小数原理来证明. 他先取一个可以被另一个数 b 量尽(即整除) 的合数 a，如果 b 是素数，那么即得结论；如果 b 不是素数，那么 b 一定可被另一数 c 量尽，那么 c 也可量尽 a，所以 c 或者是素数或者是合数. 欧几里得然后说，“如果按这种方法继续下去，那么将能找到某个素数，它可以量尽前一个数，同时也能量尽 a. 因为如果这个素数不存在，将有一个无穷数列量尽 a，而数列中的每个数都比另一个小，这对数来说是不可能的”. 这里再次看到数与量的不同，任何递减的数列总有一个最小的元，而对量来说却未必如此.

虽然欧几里得并没有这样做，但从卷 7 命题 32 很容易证明：任何数可以表示成素数的乘积. 要证明这一表示是惟一的，则需要用到卷 7 命题 30.

命题 Ⅶ – 30　如果两数相乘得某数，且一素数量尽该乘积，则它也必量尽原来的两数之一.

假设素数 p 整除 ab 但不整除 a，那么 $ab=sp$ 或 $p:a=b:s$，但是，既然 p 与 a 互素，则就这一比例而言它们是最小的数，因此 b 是 p 的倍数或 p 整除 b. 欧几里得用这一命题的结论证明了卷 9 命题 14 中素数分解的惟一性.

命题 Ⅸ – 14　如果一个数是能被一些素数量尽的最小数，那么，除了原来量尽它的素数外，任何另外的素数都量不尽它.

卷 8 主要讨论了连比例中的数，即使 $a_1:a_2=a_2:a_3=\cdots$ 的数列 $a_1,a_2,\cdots,a_n$，现代的术语称这种数列为几何数列. 现在一般认为卷 9 中的大部分内容归功于阿契塔斯(Archytas，公元前 5 世纪，他对音乐和数学感兴趣，柏拉图曾向他学习数学). 特别地，命题 8 是阿契塔斯叙述过的一个结果的推广. 原来的结论是，最终为 $(n+1):n$ 的比中的两个数没有比例中项. 想一想两根相差八度音的弦的比是 2 : 1，这个比是由 4 : 3 和 3 : 2 合成的，因此八度音是由四度和五度音合成的. 阿契塔

斯的结论是说八度音不能分成两个相等的音阶.当然,这种情况的结论等价于$\sqrt{2}$与1的不可公度性.这一结果也说明我们不能把弦长比为9∶8的全音分解成两个相同的音阶.

命题Ⅷ-8 如果在两数之间插入几个与它们成连比例的数,则无论插入多少个在它们之间成连比例的数,那么在与原来两数有同比的两数之间也能插入多少个成连比例的数.

欧几里得关注卷8中的几个其它命题,这些命题是确定在各种给定类型的数间插入比例中项数的条件.特别是命题11是卷6命题20对数的特例情形的类比.也就是说欧几里得证明了两个平方数之间有一个比例中项且两平方数之比如同它们的边与边的二次比.当然这是希波克拉底将倍立方体问题化为确定两个比例中项的问题中数的类比.

卷9是数论研究的最后一卷,该卷中的命题20是说存在无穷多个素数.

命题Ⅸ-20 预先任意给定多个素数,则有比它们更多的素数.

此命题的证明与卷6命题1的证明类似,欧几里得无法写出任意“多个”素数,因此他再次使用了“可推广举例”的方法.他只取三个素数A、B和C,再证明总可以找到另外一个素数.为此,考虑数$N = ABC + 1$,如果N是素数,那么即找到了已知三个数之外的另一素数.如果N是合数,设它可被素数p整除,接着欧几里得证明了p是与A、B和C不同的素数.这是因为A、B和C都不能整除N,因此找到了一个新的素数p.欧几里得很可能认为读者已经确信,不论开始取多少个素数,类似的证明总是可行的.

命题21-34构成了一个关于奇数和偶数的基本结论的几乎独立的体系,它们可能是毕达哥拉斯学派早期数学著作的节录,其中包括如偶数的和是偶数、奇数的偶数次的和是偶数和奇数的奇数次的和是奇数等结论.这些基本的结论可由《原本》中整个数论部分最重要的两个结论得出.

命题Ⅸ-35 如果给出成连比例的几个数,又从第二个与最后一个中减去第一个数,则从第二个数得的余数与第一个数之比等于从最后一个数得的余数与最后一个数以前各项之和的比.

事实上,这一结论相当于确定等比数列的和.用$a, ar, ar^2, ar^3, \cdots, ar^n$表示成连比例的数列,用$S_n$表示“(最后一项)前的所有数之和”(因为$ar^n$之前有$n$项),那么欧几里得的结论可以表示为

$$(ar^n - a) : S_n = (ar - a) : a,$$

这一和的现代形式是

$$S_n = \frac{a(r^n - 1)}{r - 1}.$$

卷9的最后一个命题,即命题36,讨论了怎样找到完全数,即等于其所有因子的和的数.这一结论说,如果数列$1, 2, 2^2, \cdots, 2^n$的和是素数,那么这个和与最后一项2^n的乘积是一个完全数.例如,$1 + 2 + 2^2 = 7$是素数,那么$7 \times 4 = 28$是完全数,实际上,$28 = 1 + 2 + 4 + 7 + 14$;另一个完全数是希腊人已知的6,对应于$1 + 2$;496对应于$1 + 2 + 4 + 8 + 16$;8128对应于$1 + 2 + 4 + 8 + 16 + 32 + 64$.虽然利用欧几里得判别准则已经发现了另外的一些完全数,但是否存在不符合这一准则的完全数仍属未知.特别是,迄今尚不知道是否存在奇完全数.欧几里得数论卷中最高潮的定理是研究一类在当时还只知道有四个实例的数,这也许有点奇怪.尽管如此,对数学家来说,完全数的理论始终是一个充满魅力的理论.

2.4.8 无理量

许多史学家认为卷10是《原本》中最重要的部分.它是《原本》13卷中最长的一卷,也是内容组织得最好的一卷.卷10的目的显然是对特定的不可公度量进行分类.本卷的另一个目的是想要刻

画正多面体棱长的特征，而卷 13 中正多面体的作图形成了《原本》恰到好处的顶峰. 欧几里得需要用一种非数的方法来比较二十面体和十二面体的边与它们的外接球直径的关系. 这一简单的问题引出了卷 10 的详细的分类方案，它的重要性远远超过了问题答案的本身. 既然卷 8 中一些多面体的作图被归功于蒂奥泰德，而且在柏拉图关于哪些数有不可公度的平方根的问题的对话以他为对象，那么一般认为，本卷的大部分内容应归功于蒂奥泰德.

补遗 2.6	卷 10 定义
	1. 能被同一量量尽的量叫做**可公度的量**，而不能被同一量量尽的量叫做**不可公度的量**. 2. 当一些线段上的正方形能被同一个面量尽时，这些线段叫做**正方可公度**. 当一些线段上的正方形不能被同一个面量尽时，这些线段叫做**正方不可公度**. 3. 由这些定义，我们能够证明，与已知的线段分别存在无穷多个可公度的线段与无穷多个不可公度的线段，一些仅是长度不可公度，而另外一些也是正方不可公度. 把这已知线段叫做有理线段，凡与此线段是长度、也是正方可公度或仅是正方可公度的线段，都叫做**有理线段**；而凡与此线段不可公度的线段叫做**无理线段**.

补遗 2.6 中的定义使我们能够理解欧几里得的基本术语“不可公度量(incommensurable)”和“无理线段(irrational)”的含义. 前两个定义相对明了些，但对第三个却需要作些注释. 首先，它包括一个定理，即随后在卷 10 中证明的定理. 其次，欧几里得所用的术语“有理”与现代的意义有所不同，例如，如果给定线段的长度是 1，那么不但可以把长为 a/b 的线段叫做有理线段，而且长度为 $\sqrt{a/b}$(a、b 为正整数) 的线段也可叫做有理线段.

卷 10 的第一个命题是本卷和卷 12 的基础.

命题 X－1 对于两个不相等的量，如果由较大的量中减去一个大于它一半的量，再由所得的余量中减去大于这个余量的一半的量，并且连续这样进行下去，则必得到一个余量，其小于较小的量.

这一结论是以卷 5 定义 4 为基础，即两个量有一个比的判定，这一定义要求某一个较小量的 n 倍大于较大量. 然后按步骤 n 次减大于余量的一半的量，即可得出要求的结论.

命题 2 和 3 讨论了前面已谈到的互易相减法，但是，因为欧几里得用了与卷 7 中讨论数时相同的过程，所以现在他能够把这两个概念联系起来了. 也就是说，欧几里得在命题 5 和 6 中证明了当两个量的比等于两个数的比时，那么这两个量是可公度的. 因此，虽然数和量是不同的概念，我们仍可以把数的比例理论中的方法应用于可公度的量，而欧多克斯的更复杂的定义只被用于不可公度的量.

命题 9 归功于蒂奥泰德，它推广了毕达哥拉斯学派发现的正方形的边与对角线的不可公度性，用现代的术语说即是$\sqrt{2}$ 的无理性. 实际上，欧几里得证明了每一个非平方整数的平方根相对于单位是不可公度的. 用欧几里得的原话说，即正方形的两边是长度可公度的，当且仅当正方形之间有平方数与平方数的比. 其中有趣的是“仅当”. 假设两条边 a、b 是长度可公度的，那么 $a:b=c:d$，这里 c、d 是数. 因此每个比的二次比相等. 但是欧几里得在卷 6 命题 20 中已经证明了以 a 为边的正方形和以 b 为边的正方形的比是 a 和 b 的二次比，以及 c^2 与 d^2 的比是 c 和 d 的二次比. 因此可得结论.

给出一些判别不可公度性的进一步的条件后，欧几里得开始讨论卷 10 的主要内容，无理长度的分类. 无理长度是指既不能被一个固定单位长度度量也不能正方度量的长度. 本卷的分类内容很长，因此这里只提及一些在卷 13 中应用的定义. 虽然现在可以把每个无理长度表示成一个多项式

方程的解,但欧几里得并没有用任何代数方法,所有的过程都是用几何方法.为了便于理解,这里给出每个定义的数值例子.

与由仅正方可公度的两条有理线段所构成的矩形相等的正方形的边称为**中项**(medial)线.例如,因为1和$\sqrt{5}$仅是正方可公度的,由这两个长度构成的矩形的面积等于$\sqrt{5}$,那么长度$\sqrt[4]{5}$是中项.仅正方可公度的两有理线段之和称为**二项**(binomial)线,因此长为$1+\sqrt{5}$的线段是二项线.类似地,仅正方可公度的两有理线段之差称之为**余线**(apotome).长为$\sqrt{5}-1$的线段便是余线.最后,一个更为复杂的例子是欧几里得关于**次线**(minor)的定义.两条线段x和y正方不可公度,且使得x^2+y^2是有理的和xy是中项面(即等于一中项线上的正方形),那么$x-y$便称为次线.例如,如果$x=\sqrt{5+2\sqrt{5}}$、$y=\sqrt{5-2\sqrt{5}}$,那么$x-y$便是次线.

2.4.9 立体几何

《原本》共有三卷讨论立体几何,卷11是其中的第一卷.它包括了与卷1和卷6中许多二维结论相似的三维结论,引入的定义有棱锥、棱柱和圆锥(补遗2.7).球的定义有些特殊,它不是用类似圆的定义给出的,而是通过以直径为轴旋转来定义的.这可能是因为欧几里得不想像在卷3中讨论圆的性质那样来讨论球的性质.实际上,在欧几里得时期,球的基本性质已为大家所知并被记载到其它的著作中,其中包括欧几里得本人的另一本著作.但是,欧几里得只在《原本》卷12和卷13中讨论了球:卷12考察了球的体积,卷13给出了正多面体的作图并说明了如何把它们放入球中.他在卷13中的作法如他对球的定义一样,实际上是通过旋转一个半圆来使这些多面体内接于一个球.

	卷11中的定义
补遗 2.7	12.一个**棱锥**是一个立体图形,它是由一些面构成,包括一个给定的面及由此面到一定点作成的面. 13.一个**棱柱**是一个立体图形,它是由一些平面构成的,其中有两个面是相对的、相等的、相似且平行的,其它各面都是平行四边形. 14.固定一个半圆的直径,旋转半圆到开始位置所形成的图形是一个**球**. 18.固定直角三角形的一条直角边,旋转直角三角形到开始位置所形成的图形是一个**圆锥**.如果所固定的直角边等于另一直角边,这时所形成的图形是**直角圆锥**;如果小于另一边,则是**钝角圆锥**;如果大于另一边,则是**锐角圆锥**.

卷11中的命题包括一些与卷1类似的作图问题.例如,命题11给出了怎样过平面外一点作平面的垂线,命题12是怎样过平面上的一点作平面的垂线.也有一些关于平行六面体的定理.特别地,与卷1命题36相似,欧几里得证明了同底等高的平行六面体相等,与卷6命题1相似,他也证明了同高的平行六面体的比等于它们底的比.同样与卷6命题19和20相似,他在命题33中证明了相似平行六面体的比是它们边的三次比,亦即相似平行六面体的体积比等于其对应边的立方比.与前面一样,欧几里得没有计算体积,但从这些定理可以得出平行六面体体积的基本结果.卷12给出了其它立体的体积"公式".

卷12与其它卷不同的特征是它使用了极限过程,一般以穷竭法著称.这方法是由欧多克斯发展并用于计算圆的面积和棱锥、棱柱、圆锥及球的体积.这些面积和体积"公式"中有一些更早之前就已经被发现了,但希腊数学家觉得还必须加以证明,而欧多克斯的方法给出了这样的证明.当然

这里并没有给出最初发现这些公式的方法.

卷 12 中其它的结论如下:

命题 Ⅻ – 2 圆与圆之比等于直径上的正方形之比.

命题 Ⅻ – 7 (推论)任何一个以三角形为底的棱柱可以被分成以三角形为底的三个彼此相等的棱锥.

命题 Ⅻ – 10 圆锥是与它同底等高的圆柱的三分之一.

命题 Ⅻ – 18 球的比等于它们直径的三次比.

上面的第一个结论是关于圆面积结果的欧几里得形式,150 年前的希波克拉底已经发现了这一形式.用现代术语来说,即圆的面积与直径的平方成比例.它并没有说比例常数是什么,但证明过程给出了其近似计算方法.卷 12 命题 1 是作为上述证明的一个引理,该命题是说圆内接相似多边形的比等于圆直径上的正方形之比.这一结果本身又是卷 6 命题 20,即相似多边形的比等于它们对应边的二次比的推广.首先不难证明,可以用任意的对应线段代替"对应边",包括圆的直径,其次容易明了,可以用"平方"代替"二次比".

卷 12 命题 2 的证明思想主要是用边数不断增加的内接多边形来"穷竭"一个圆.特别地,欧几里得指出,可以在一个已知圆中内接一个多边形,其面积与圆面积之差小于任意给定的数.他的证明是先假设结论不正确,也就是说,如果两个圆 C_1 和 C_2 的面积分别为 A_1 和 A_2,直径分别为 d_1 和 d_2,设 $A_1 : A_2 \neq d_1^2 : d_2^2$,那么将有一个面积 S,或者大于或者小于 A_2,使得 $d_1^2 : d_2^2 = A_1 : S$.首先假设 $S < A_2$(图 2.36),那么从内接正方形开始,依次二等分所对的弧,在 C_2 中内接一个多边形 P_2 且 $A_2 > P_2 > S$,也就是 P_2 与 A_2 的差小于 A_2 与 S 的差,这种作图可能是根据卷 10 命题 1 进行的,因为每次二等分弧都使得多边形面积的增加大于圆与多边形面积差的一半.接下来在 C_1 中内接一个与 P_2 相似的多边形 P_1.由卷 12 的命题 1 可知,$d_1^2 : d_2^2 = P_1 : P_2$,根据前面的假设,那么这一比也等于 $A_1 : S$,所以 $P_1 : A_1 = P_2 : S$,但显然 $A_1 > P_1$.由此可得 $S > P_2$,与假设 $P_2 > S$ 矛盾,因此 S 不能小于 P_2.通过归结为已经说明了的情形,欧几里得同样证明了 S 也不能大于 A_2.所以正如所要证明的那样,圆的比一定等于其直径上的正方形之比.

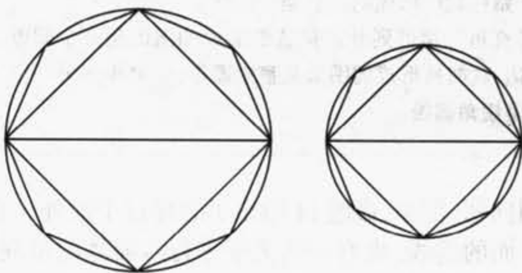


图 2.36 《原本》卷 12 命题 2:穷竭法.



图 2.37 希腊邮票上的德谟克里特.

埃及人和巴比伦人一定知道计算棱锥体积的定理(补遗 2.8).但是阿基米德写道,德谟克里特(Democritus,公元前 5 世纪)第一个发现了这一结论(图 2.37),欧多克斯第一个给出它的证明.遗憾的是我们没有关于埃及、巴比伦人和德谟克里特是怎样发现这些结论的记载.克里斯帕斯的一篇报导记载了德谟克里特讨论的问题,即用平行于圆锥底面的平面把圆锥截为"不可分"的部分,德谟

克里特想知道这些不可分的圆究竟是相互相等还是不相等"若它们不相等,那么圆锥是不规则的,如像阶梯那样,有许多锯口和不均匀;如果它们相等,那么截面是相等的,圆锥将有圆柱的性质,由相等的而不是不相等的圆组成,这是很荒唐的."[22]

补遗 2.8

希腊人从埃及学到了什么?

希腊人学习了埃及数学吗?希腊的数学思想与其前人有所不同,这使我们猜想他们的研究是从新的起点开始的.这一问题已提出了很久,但因为没有3世纪前埃及向希腊传播数学的有关文献,所以我们不能给出确切的答案,但仍有一点线索.

一般来说,希腊人承认曾向埃及学习.许多希腊数学家如毕达哥拉斯、泰勒斯和欧多克斯等的故事都讲到了他们曾在埃及学习过.一些希腊的文献中也说埃及人首先发明了几何学然后传到希腊.但是"几何学"的意思是什么?它肯定不是指我们所知的如欧几里得《原本》中的公理化处理.它可能指的是结论本身.人们毕竟不能靠公理方法来发现结论,而是通过实验、试错或归纳来发现.只有在作出发现以后,人们才开始考虑这些结论是否正确.那么希腊文献中所说的埃及人发明了几何应该是指结论而非证明方法.显然证明的公理体系起源于希腊.

希腊人向埃及学了些什么几何结论?一个答案可能是与几何对象测量有关的大多数公式,如棱锥、圆的面积和半球的面积等.他们也可能学习了相似的基本原理,因为从埃及的文献中我们知道,埃及已经有了与标尺使用有关的较先进的比例思想.我们还可以确定,希腊人从埃及学了单位分数的应用,虽然在希腊正式的数学著作中这些知识并没有出现.

和巴比伦的情况一样,我们没有埃及影响希腊数学的直接文献证据,但间接的证据还是相当充分的.因此我们期待进一步的研究以给出这些答案.[21]

虽然我们不知道德谟克里特最后的结论是什么,但他肯定考虑到圆锥以及类似地棱锥是由不可分量"组成".如果那样的话,他可能已经得到了欧几里得卷12中的命题5,即以三角形为底且同高的两个棱锥的比等于两底的比.因为如果假设用一与底面平行且到底面等距的平面分别截两棱锥,那么对应截面的比将等于底的比,既然德谟克里特想到每个棱锥可看作是由无数多个不可分的截面"组成",那么棱锥之间的比就等于对应截面的比.然而他可应用与《原本》卷12命题7相似的结论(即底为三角形的棱柱可以分解为三个等高等底的棱锥)来完成证明.

当然,欧几里得用反证法证明了卷12的命题5、10和18.假设给出的结论不正确,接着他在一给定的立体中作另一些性质已知的立体,使得它们与给定立体的差小于任一给定"小"的值,即由不正确的假设导出的"误差".于是他穷竭了这个立体.然后像卷7命题2的证明那样,由所作图形的已知性质引出矛盾.虽然亚里士多德在希腊数学著作中禁止使用无穷小的概念,但我们从所引用的德谟克里特的内容中可知,从希腊数学的早期开始,一些数学家们就打算借助无穷小来发现一些结论.

《原本》的最后一卷——卷13讨论了五个正多面体作图以及它们"包含"在一个球中的情况(图2.38).此卷内容是卷4的三维类似.五个正多面体——立方体、四面体、八面体、十二面体和二十面体的研究以及只存在这五个正多面体的证明应归功于蒂奥泰德.在希腊时期之前,人们已经知道了前三种立体,公元前7世纪的青铜制正十二面体也已被发现,但正二十面体显然是首先由蒂奥泰德所研究,也正是他注意到只存在这五种正多面体,并且认为正多面体的性质值得探讨.

欧几里得在卷13中系统地研究了每个正多面体的作图,证明了每个都可以包含(内接)在一个球中,还把它们的棱长与球的直径作了比较.他证明了,对正四面体,直径上的正方形是棱上正方形

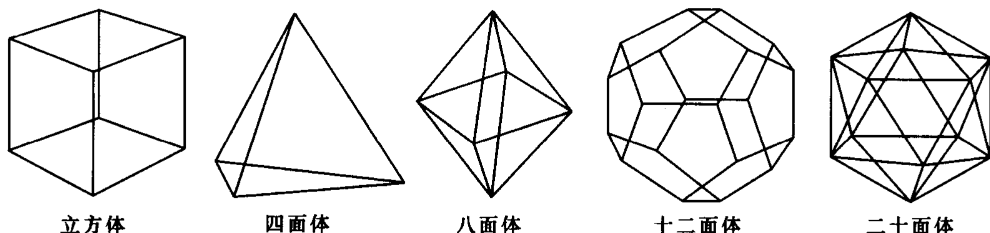


图 2.38　五个正多面体:立方体、四面体、八面体、十二面体和二十面体.

的$1\frac{1}{2}$倍;对立方体,则是3倍;对正八面体,则是2倍.对另外两个立体的情况,则需要一些技巧.欧几里得证明了正十二面体的棱是长度等于当一个内接立方体的棱按中外比分割时较长的线段的余线.因此,如果球的直径是1,那么立方体的棱长$c=\sqrt{3}/3$,所以正十二面体的棱长等于方程$x^2+cx-c^2=0$的正根,或$(c/2)(\sqrt{5}-1)=(1/6)(\sqrt{15}-3)$.因为按欧几里得的定义,$\sqrt{15}$和$\sqrt{3}$是有理的,又因为它们仅仅是正方可公度的,所以边长事实上确是余线.

对二十面体,欧几里得证明了棱是次线.此时,球的直径上的正方形是二十面体上部五个三角形的外接圆半径上正方形的5倍.这五个三角形的底构成一个正五边形,其边是二十面体的棱,这个内接于半径为r的圆的五边形的边长等于

$$\frac{r}{2}\sqrt{5+2\sqrt{5}}-\frac{r}{2}\sqrt{5-2\sqrt{5}}=\frac{r}{2}\sqrt{10-2\sqrt{5}}.$$

如果球的直径是1,那么r是有理值且等于$\sqrt{5}/5$,二十面体的棱长的确是一条次线,且它的长度等于

$$\frac{\sqrt{5}}{10}\sqrt{10-2\sqrt{5}}=\frac{1}{10}\sqrt{50-10\sqrt{5}}.$$

欧几里得在一个平面图形中作出了这五个正多面体的棱,同时还把它们与已知球的直径作了比较,然后他证明了除了这五个外再没有其它的正多面体.这是卷13和整个《原本》圆满的终曲.

2.5　欧几里得的其它著作

欧几里得写了几本比《原本》更高级的书,其中现存最重要的一本是《数据(Date)》[23],它实际上是《原本》卷1-6的补充.《数据》中的每个命题都是已知几何图形内的某些元素,证明可以确定另外某些元素.它实质上是把《原本》中纯粹综合的内容改编为手册的形式以适合希腊数学的宗旨之一:解决新问题.

这里只给出该书中的两个命题,它们与《原本》卷6命题29密切相关,也研究几何代数.

命题 84　如果两条线段以一定的夹角构成一个已知面积,且两线段的差已知,那么这两线段即为已知(即可确定).

正如《原本》卷6命题29的讨论一样,如果假设已知角是直角(现存的中世纪手稿中给出了直角情况的图形),那么这与一个标准的巴比伦问题相关:已知x、y的积与差,求x、y,即解联立方程

$$xy=c,\quad x-y=b.$$

欧几里得作一个边分别为x、y的矩形,在x上减去y.现在他将一个已知面积C贴合到已知线

段 b 上，且外加一个正方形，然后他应用命题 59.

命题 59　如果将一已知面积贴合到一条已知线段上，且外加一个大小已知的图形（即其角以及边的比为已知的图形），那么可以确定这个外加面积的边.

这里，欧几里得使用与卷 6 命题 29 相同的图解决了命题 84 的问题. 他把长度 b 二等分，再在 $b/2$ 上作正方形，而这个正方形与面积 c 的和等于在 $y + b/2$（或 $x - b/2$）上的正方形，因此给出了确定作为正方形的边的这些量的方法. 若用代数形式表示，这就相当于标准的巴比伦公式

$$y = \sqrt{\left(\frac{b}{2}\right)^2 + c} - \frac{b}{2}, \qquad x = \sqrt{\left(\frac{b}{2}\right)^2 + c} + \frac{b}{2}.$$

与前面一样，欧几里得只讨论几何图形，而从不写出这样的计算法则. 但我们可以很容易地把他的几何形式转换为代数形式. 上面的问题实际上是确定满足特定条件的两条线段的长度，也就说即使是问题的陈述形式也几乎与巴比伦人是一样的. 另一方面，如卷 6 命题 29 那样，这个结果使我们不仅能像巴比伦人那样讨论矩形情形，而且也能处理平行四边形情形. 欧几里得在《数据》中还讨论了类似的其它几何代数问题，如命题 85 和 58 中与 $xy = c$、$x + y = b$ 等价的几何问题.

欧几里得的著作还涉及光学、音乐和圆锥曲线等. 从他的这些著作来看，他好像把自己当作一个到他那个时代的希腊数学的汇编者. 如果他是第一个被邀请到亚里山大的博物院（Museum）的数学家，那么这一判断应该是正确的. 因此欧几里得的目的不只是要向学生们传授当时已经发现的结论，同时还要向他们传授发现新问题的方法. 公元前 3 世纪的两位数学家阿基米德和阿波罗尼乌斯极大地拓宽了数学的研究领域，很可能他们接受的最早的数学教育就是来自欧几里得的学生，因此他们能够解决欧几里得及其学生留下的未解决的许多问题.

习　题

泰勒斯的问题

1. 据说泰勒斯曾给出一种应用"角边角"定理计算船离河岸距离的方法，一种可能的方法是：设 A 是河岸上一点，S 是船（图 2.39），沿 AS 的垂线测量距离 AC，取 AC 的中点 B，过点 C 作一与 AC 垂直的直线，在此直线上取一点 E，使得 E、B、S 共线，试证 $\triangle EBC \cong \triangle SBA$，因此 $SA = EC$.

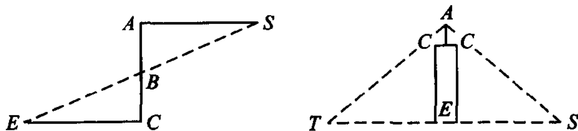


图 2.39

2. 泰勒斯的第二种可能的方法是：假定泰勒斯站在河岸的塔顶上，拿着一个由一根直棒和一横木 AC 做成的装置，这个装置可旋转任意角度后行止不动. 那么转动 AC，使得沿 AC 方向可以看到船 S，然后再转动 AC 使得可以看到岸上的某物 T，试证 $\triangle AET \cong \triangle AES$，因此 $SE = ET$.

毕达哥拉斯学派的数论问题

3. 试证第 n 个三角形数的代数表示形式为 $T_n = \frac{n(n+1)}{2}$,因此一个矩形数等于一个三角形数的两倍.
4. 用代数方法证明一个正方形数是两个连续三角形数之和.
5. 用画点法来证明任意三角形数的8倍加1是一个正方形数;任意奇正方形数减1是一个三角形数的8倍.再用代数方法证明以上结论.
6. 试证在毕达哥拉斯三元数组中,如果其中一项是奇数,那么另外两项必定一奇一偶.
7. 试证在毕达哥拉斯三元数组中,如果最大的一项可被4整除,那么另外两项也可被4整除.
8. 叙述并证明在第7题中用"3"代替"4"的情况.
9. 当 n 是奇数时,用公式 $\left(n,\frac{n^2-1}{2},\frac{n^2+1}{2}\right)$ 作出五个毕达哥拉斯三元数组,当 m 是偶数时,用 $\left(m,\left(\frac{m}{2}\right)^2-1,\left(\frac{m}{2}\right)^2+1\right)$ 作出另外不同的五个毕达哥拉斯三元数组.
10. 用类似于毕达哥拉斯学派证明 $\sqrt{2}$ 和1是不可公度的方法证明 $\sqrt{3}$ 和1也是不可公度的.

欧几里得《原本》中的问题

11. 试证命题Ⅰ-32:任意三角形的三个内角和等于两个直角.说明这一证明以Ⅰ-29为基础,因而也以公设5为基础.
12. 解命题Ⅰ-44的问题(略有修改):在已知线段 AB 上贴合一与已知矩形 c 相等的矩形.如图2.40,其中 $BEFG$ 是已知矩形,D 是对角线 HB 的延长线与 FE 延长线的交点,$ABML$ 就是要作的矩形.

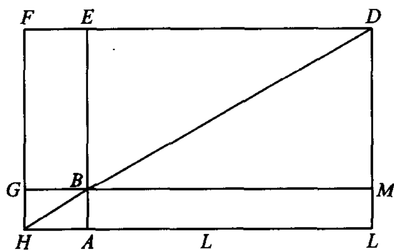


图2.40　《原本》命题Ⅰ-44.

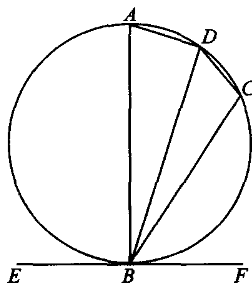


图2.41　《原本》命题Ⅲ-32.

13. 把命题Ⅱ-8改写成代数结论,并证明它是正确的:如果任意两分一条线段,用原线段和所分得的较小线段构成的矩形的4倍与分得的另一线段上正方形的和等于原线段与较小线段之和上的正方形.
14. 试证命题Ⅱ-13(等价于锐角三角形的余弦定理):锐角三角形中,锐角所对边上的正方形比另外两边上正方形的和小一个矩形的2倍,该矩形是由另一锐角向对边作垂线,垂足到原锐角之间的一段与该边所构成.(命题Ⅱ-12给出了钝角三角形的余弦定理.)
15. 给出命题Ⅲ-20的详细证明:在一个圆中,同弧上的圆心角等于圆周角的2倍.
16. 给出命题Ⅲ-21的详细证明:在一个圆中,同弧上的角彼此相等.
17. 试证命题Ⅲ-31:半圆上的角是直角.
18. 试证命题Ⅲ-32:如果一条直线切于一个圆,且由切点作一条过圆内部的直线和圆相截,那么该直线与切线所成的角等于相应截弧上的角.如图2.41,EBF 是切线,BD 是割线,AB 是直径,C 是弧 DB 上的任意点,结论即是 $\angle FBD=\angle BAD$,$\angle EBD=\angle DCB$.利用 $\angle ADB$ 是直角这一事实.

19. 给出作圆内接正六边形的方法.

20. 已知正五边形和正三角形可以内接于一个圆,说明怎样作一个圆内接正 15 边形.

21. 证明在已知数 a、b 的欧几里得算法中,最后一个非 0 的余数是 a、b 的最大公因子.

22. 用欧几里得算法分别求 963 和 657、2689 和 4001 的最大公因子.

23. 假设把长度为 1 的一条线段分成中外比,即按 x 分线段,使得 $1/x = x/(1-x)$. 用欧几里得算法证明 1 和 x 是不可公度的,也就是证明可以用蒂奥泰德的定义把 $1 : x$ 表示成$(1,1,1,\cdots)$ 的形式.

24. 用蒂奥泰德等比的定义证明 $46 : 6 = 23 : 3$,这两个比都可以表示成$(7,1,2)$ 的形式.

25. 根据所涉及的量的类型及所给的证明,比较欧几里得在命题 Ⅴ-1 和命题 Ⅱ-1 中对分配律的处理.

26. 用欧多克斯的定义和现代的方法证明命题 Ⅴ-12:如果有任意个量成比例,那么其中一个前项比相应的后项等于所有前项的和比所有后项的和.(用代数符号表示,即如果 $a_1 : b_1 = a_2 : b_2 = \cdots = a_n : b_n$,那么

$$(a_1 + a_2 + \cdots + a_n) : (b_1 + b_2 + \cdots + b_n) = a_1 : b_1.)$$

27. 用欧多克斯的定义证明命题 Ⅴ-16:如果 $a : b = c : d$,那么 $a : c = b : d$.

28. 用几何方法构作 $8 : 4 = 6 : x$ 的解.

29. 已知半径为 $9 + 5 = 14$ 的圆,用几何方法解方程 $9/x = x/5$.

30. 用几何方法(Ⅵ-28) 解二次方程 $x^2 + 10 = 7x$. 此方程有两个正数解. 通过变化图形,使得这两个解变得显然.

31. 用几何方法(Ⅵ-29) 解二次方程 $x^2 + 10x = 39$.

32. 证明命题 Ⅷ-8 和阿契塔斯的特例:$(n+1)$ 与 n 之间没有比例中项.

33. 确定命题 Ⅷ-11 中两个平方数的比例中项.

34. 确定命题 Ⅷ-12 中两个立方数的比例中项.

35. 试证命题 Ⅷ-14:如果 a^2 量尽 b^2,那么 a 也量尽 b. 反之亦成立.

36. 用命题 Ⅶ-30 证明任意正整数素数分解的惟一性(不计次序)(即命题 Ⅸ-14).

37. 把欧几里得 Ⅻ-2 的证明变为计算圆面积的递推算法,并通过几次计算,求出半径为 1 的圆面积的近似值.

38. 试证命题 ⅩⅢ-9:如果将内接于同一个圆的正六边形的一边与正十边形的一边加在一起,则它们的接点将将此和线段分为中外比,其中较长线段是正六边形的边. 如图 2.42, BC、CD 分别是内接于同一个圆中的正十边形、正六边形的边,证明 $\triangle EBD$ 与 $\triangle EBC$ 相似.

图 2.42 《原本》命题 ⅩⅢ-9.

39. 证明命题 ⅩⅢ-10:如果有一个内接于圆的正五边形,那么其一边上的正方形等于同圆的内接正六边形一边上的正方形与内接正十边形一边上的正方形的和. 取半径为 1 的圆,用已知的内接正多边形的边长的值加以验证.

40. 用《数据》中的几何方法解二元方程 $xy = 7$, $x - y = 18$.

讨论题

41. 埃拉托塞尼(Eratosthenes of Cyrene,公元前 276—194) 曾用平行线的方法测量过地球. 他发现,在夏至中午,太阳在悉恩(Syene,即现在的阿斯旺,埃及东南部,尼罗河河畔都市,位于北回归线上) 的正上方,而在距北回归线 5000(stades) 的亚里山大,这时太阳偏离正上方 $7\frac{1}{5}°$. 设光线从太阳到地球是平行的,埃拉托塞尼断言 $\angle SOA = 7\frac{1}{5}°$(图 2.43). 那么,计算埃拉托塞尼的地球周长的值(以 stades 为单位). 如果取 1 stades 的长等于 516.7 英尺(= 300 埃及肘尺),计算埃拉托塞尼的地球周长和直径的值. 这种方法的精确度如何?埃拉托塞尼是如何知道亚里山大到悉恩的距离的?

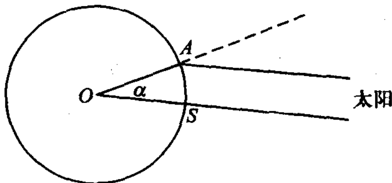


图 2.43　埃拉托塞尼测定地球.

42. 希腊人通过把巴比伦的代数方法译成几何形式,从而发展了它的几

何代数.试简述赞成或反对这一观点的理由.

43.试讨论:在二次方程的教学中,几何方法相对于纯代数方法的优缺点.

44.准备一节用几何方法证明简单的代数恒等式的课.(例如,证明$(a+b)^2=a^2+2ab+b^2$和$(a+b)(a-b)=a^2-b^2$.)

45.比较欧几里得判断三角形全等的方法和现在中学几何教材中的方法,哪种方法更容易讲授?

46.试讨论:欧几里得的《原本》是否符合柏拉图的格言——几何的研究是为了"把灵魂引向真理",为了获得"永恒"知识.

47.试讨论柏拉图在《理想国》中所说的算术和几何之类的研究在军事上的应用.一个将领必须是数学专家吗?为什么柏拉图提到的数学研究惟一的"实际"应用是军事?

48.高中的几何学习是否应以多年的欧几里得《原本》学习为基础?讨论欧几里得的方法与"现代"方法相比较的利与弊.[24]

文献和注解

了解希腊文明基本情况的较好著作是 H.D.F.Kitto 的 *The Greeks*(London:Penguin,1951).而对早期希腊科学论述较好的两部著作分别是 G.E.R.Lloyd 的 *Early Greek Science: Thales to Aristotle*(New York:Norton,1970)和 *Magic, Reason and Experience*(Cambridge:Cambridge University Press,1979),特别地,后者主要研究了逻辑推理在希腊的产生及数学证明思想的出现.了解希腊数学的常用的参考书籍是 Thomas Heath 的 *A History of Greek Mathematics*(New York:Dover,1981,reprinted from the 1921 original),但是,Heath 的许多论断已经受到了以下及随后三章中所列的更近代的研究的挑战.Heath 还出版了欧几里得《原本》的标准的现代英文版本(Euclid's *Elements*(New York:Dover,1956)).Ivor Thomas 的 *Sellections Illustrating the History of Greek Mathematics*(Cambridge:Harvard University Press,1941)中收集了早期希腊数学的原始资料.B.L.Van der Waerden's 的 *Science Awakening*(Groningen:Noordhoff,1954)一书中也包含对希腊数学的广泛研究.Wilbur Knorr在其 *The Evolution of the Euclidean Elements*(Dordrecht:Reidel,1975)中对有关欧几里得《原本》的产生背景给出了更为详尽但多少引起争议讨论.Knorr 一部更新的著作 *The Ancient Traditional of Geometric Problems*(Boston:Birkhäuser,1986)则充分地论证了求解几何问题是促进希腊数学发展的主要原因. 另一本重构希腊数学背景的著作是 D.H.Fowler 的 *The Mathematics of Plato's Academy: A NewReconstruction*(Oxford:Clarendon Press,1987)该书认为互易相减(*anthyphairesis*)的思想为希腊数学提供了重要动力.本章在讨论欧几里得思想时也引用了 Charles Jones 的博士论文 *On the Concept of One as a Number*(University of Toronto,1979)和 I.Mueller 的博士论文 *Philosophy and Deductive Structure in Euclid's Elements*(Cambridge:MIT Press,1981)中的资料.关于希腊数学的其它有意义的著作还有:F.Lasserre 的 *The Birth of Mathematics in the Age of Plato*(Larchmont,N.Y:American Research Council,1964),J.Klein 的 *Greek Mathematical Thought and the Origin of Algebra*(Cambridge:MIT Press,1968)和 Asger Aaboe 的 *Episodes from the Early History of Mathematics*(Washington:MAA,1964).最后,在 J.L.Berggren 的"History of Greek Mathematics:A Survey of Recent Research,"(*Historia Mathematica* 11(1984),394—410)中给出了近来对希腊数学研究的文献综述.

1.引自 Proclus' Summary,英译文见 Thomas, *Selections Illustrating*.I,p.147.

2.Phillip H.De Lang 和 Benedict Einarson 译, *Plutarch's Moralia*,(Cambridge:Harvard University Press,1959),Ⅶ,pp.397—399.

3.Plato, *Republic* VII,525.这里采用的是 Frances Cornford 的英译本(London:Oxford University Press,1941),但参照了标准的希腊文本,可与现代的任一译本对照.

4.同上,VII,527.

5.同上,VII,528.

6.同上,VII,531.

7.Aristotle, *Prior Analytics* I,1,24^b,19.这里所引亚里士多德著作的译文是根据 Great Books 版(Chicago:Encyclopedia

Britannica,1952),但像柏拉图著作的情形一样,也参照了标准的希腊版本.

8. Aristotle, *Posterior Analytics* I.10,76ᵃ 40 - 76ᵇ 10.

9. 同上,I.2,71ᵇ 23.

10. Aristotle, *Physics* VI,1,231ᵇ 15.

11. 同上,V,3,227ᵃ 12.

12. 同上,VI,9,239ᵇ 11. 有关芝诺悖论的更多资料参见 F.Cajori,"History of Zeno's Arguments on Motion,"*American Mathematical Monthly* 22(1915).1-6,39-47,77-82,109-115,145-149,179-186,215-220,253-258,292-297,或 H.D.P.Lee, *Zeno of Elea*(Cambridge:Cambridge University Press,1936).

13. 同上,Ⅵ,9,239ᵇ 15.

14. 同上,Ⅵ,2.233ᵇ 26 - 29.

15. 同上,Ⅵ,9,239ᵇ 6.

16. Proclus, *Commentary on the first Book of Euclid's Elenments*, (G.P.Morrow 译,Princeton:Princeton University Press, 1970),p.56.此版本中也包含了一个对《原本》有价值的介绍.

17. Sabetai Unguru 在题为"On the need to rewrite the history of Greek mathematics,"(*Archive for History of Exact Sciences* 15(1975),67—114)的文章中,重新挑起了关于几何代数的猛烈争论.随后两年里,其他的几位数学史家对此作了回应,其中最重要的要数 B.L.Van der Waerden 的"Defence of a shocking point of view,"(*Archive for History of Exact Sciences* 15(1976),199 - 210,)和 Hans Freudenthal 的"What is algebra and what has it been in history?"(*Archive for History of Exact Sciences*,16(1977),189 - 200).Unguru 和 David Rowe 在"Does the Quadratic Equation Have Greek Roots?A Study of Geometric Algebra,Application of Areas,and Related Problems,"(*Libertas Mathematica* 1(1981)和2(1982))中对此又作出了回应.这些文章可以作为历史争论所激起的感情冲动的很好的例子.

18. 有关中外比理论的更多详细内容参见 Roger Herz-Fischler, *A Mathematical History of Division in Extreme and Mean Ratio*(Waterloo, Ont.:Wilfrid Laurier University Press,1987).

19. D.H.Fowler,The Mathematics of Plato 's Academy,p.225.其中对在希腊数学发展过程中辗转相除(anthyphairesis)及其可能的意义给出了详细讨论.也可参见 D.H.Fowler,"Ratio in Early Greek Mathematics,"*Bulletin of the American Mathematical Society* 1(1979),807 - 847.

20. 关于欧多克斯比例理论产生的讨论引自 Knorr 的 *The Evolution of the Euclidean Elements*.

21. 近年来,史学界对希腊文明和埃及文明之间的关系有些争论,特别是希腊数学和埃及数学.公开论战的起点是 Martin Bernal 的 *Black Athena:The Afroasiatic Roots of Classical Civilization*(New Brunswick:Rutgers University Press, 1987)之出版,书中断言,古典希腊文明具有深刻的亚非文化根源,但自18世纪以来,主要由于种族原因,这种影响完全被忽视或否定.Bernal 在此书中对科学本身并没有很多的论述,但在"Animadversions on the Origins of Western Science."(*Isis* 83(1992),596 - 607)中综述了他的关于埃及科学对希腊科学的贡献的观点.随后,Robet Palter 在"*Black Athena*,Afro-Centrism,and the History of Science,"(*History of Science* 31(1993),227 - 287)一文中对此观点提出质疑,Bernal 又在"Response to Robert Palter,"*History of Science* 32(1994),445 - 464)中作了回答,而 Palter 在同期杂志上(pp.464 - 468)回复了 Bernal.关于此问题的争执尚无定论.

22. Thomas,Selections Illustrating,I,p.229.

23. 欧几里得的《数据》(*Data*)有两个英文版本,一个是 Shuntaro Ito 的 *The Medieval Latin Translation of the Data of Euclid*(Boston:Birkhäuser,1980),另一个是 George L.McDowell 和 Merle A.Sokolik 的 *The Data of Euclid*(Baltimore: Union Square Press,1993).

24. 参见 Paul Daus,"Why and How We Should Correct the Mistakes in Euclid,"*Mathematics Teacher*53(1960),576 - 581.

公元前300年之前的希腊数学概览

年代	人物/事件	成就
公元前624—547 *	泰勒斯(Thales)	定理的“证明”
公元前572—497	毕达哥拉斯(Pythagoras)	万物皆数
公元前5世纪	希波克拉底(Hippocrates of Chios)	新月形、倍立方体
公元前5世纪	芝诺(Zeno)	运动悖论
公元前5世纪	阿契塔斯(Archytas)	音乐和数论
公元前430	正方形的边与对角线是不可公度的	
400 B.C.	狄奥多西(Theodorus)	不可公度
公元前429—347	柏拉图(Plato)	385年建立学园
公元前417—369	蒂奥泰德(Theaetetus)	不可公度量、比例理论
公元前408—355	欧多克斯(Eudoxus)	比例理论
公元前384—322	亚里士多德(Aristotle)	三段论逻辑
300 B.C.	建立了亚里山大博物院和图书馆	
300 B.C.	欧几里得(Euclid)	《原本》
公元前280—206	克里斯帕斯(Chrysippus)	命题的逻辑性

* 本表中所列均为公元前年份 —— 译注.

第3章 阿基米德和阿波罗尼乌斯

> 《圆锥曲线》第三卷包括许多对综合研究立体曲线(solid loci)很有用的重要定理……;这里绝大多数最优美的定理都是新的,这些定理的发现使我知道欧几里得并没有综合地解决三线和四线轨迹问题……;因为没有我发现的这些定理作基础,不可能完成对这些问题的所谓综合.
>
> (阿波罗尼乌斯《圆锥曲线》卷1序)[1]

维特鲁维乌斯(Vitruvius)曾讲过这一故事,"在他(希罗(Hiero),公元前3世纪叙拉古的国王)取胜后,决定在寺庙里铸造一个金王冠,以献给永恒的神.他对工匠的报酬及铸造王冠所需的黄金量都作了详细的计算.工匠在指定时间内做好了王冠,并送交国王验收.王冠的重量看起来完全符合国王的要求,不过有人透漏说,在王冠制作过程中,有一部分黄金被用相等重量的白银替换了.希罗王知道后很生气,认为自己被欺骗了,但他又没办法证明黄金是否被白银所代替,于是他便请教阿基米德.阿基米德一直考虑着这一问题,但是想不出好办法.碰巧有一次,阿基米德在洗澡时发现,溢出到浴池外面的水的重量与他浸在水中的那部分身体的重量相等,于是阿基米德想出了解决王冠问题的方法.他马上跳出浴池,向家里跑去,而且边跑边喊:"我找到了!我找到了!"[2]

公元前3世纪到公元前2世纪初,希腊数学的两个主要代表人物是叙拉古的阿基米德(287 B.C.—212 B.C.)和波格的阿波罗尼乌斯(250 B.C.—175 B.C.),他们分别继承了公元前4世纪希腊数学的不同风格.阿基米德继承了欧多克斯的"极限"方法,他不仅把这一方法应用于新的面积和体积的计算,而且还给出了新的计算技巧,同时用这些技巧得到了许多新的结果.阿基米德与其前人不同,他既不勉强别人接受其发现的方法,也不厌烦数值计算并给出数值结果.他也写了我们今天所称之为理论物理的一些数学模型的几部著作;他还把他的物理原理应用到许多机械发明中.

阿波罗尼乌斯则把分析方法推广应用到新的更复杂的几何作图问题.他这些成就的基础是其8卷本的巨著《圆锥曲线论》.书中阿波罗尼乌斯系统地给出了这类曲线的重要性质,而这些性质在他得到倍立方体、三等分角等问题的新结论的过程中占有核心地位.

本章综述了这两位数学家现存的著作和研究类似问题的其他一些人的著作.

3.1　阿基米德和物理学



图 3.1　阿基米德及杠杆原理.

阿基米德是通过对物理问题构造数学模型而得到定量结果的第一位数学家.特别地,他首次证明了杠杆原理,并应用这一原理找到了物体和图形的重心;同时他还首次证明了流体静力学原理,并给出了这一原理的重要应用.

3.1.1　杠杆原理

在小时候玩跷跷板时,大家就已经熟悉了杠杆原理:如果距支点两端等距离的重量相等,那么杠杆将保持平衡.所以一个较轻的小孩可以通过远离支点来和一个较重的小孩达到平衡状态.古人也充分认识到了这一点,甚至早在亚里士多德的一本手抄的机械著作中就已经出现了这一原理:"两个重量相等的物体,离支点远的要比离支点近的运动得更快,在杠杆上有三个元素,支点 …… 和两个物体,其中一个物体动,另一个也应该跟着动,而两物体重量之比与它们到支点的距离成反比."[3]

据了解,在阿基米德之前没有人建立过杠杆的数学模型,而用此模型可以给出杠杆原理的数学证明.一般地,把数学应用于解决物理问题的一个困难往往是在于物理状态的复杂性.这就需要把这个状态理想化,所以我们总是忽略问题中的一些看上去不很重要的方面,而只集中考虑物理问题中的基本变量,这种理想化过程我们今天仍称之为建立数学模型.杠杆就是一个相关的例子.

人物小传

阿基米德(287 B.C.—212 B.C.)

在古希腊的众多数学家中,后人对阿基米德的传记性描述最多.在第二次布匿战争中,罗马将军马塞勒斯于公元前 212 年攻陷了西西里的主要城市叙拉古.普卢塔克为马塞勒斯作的传记中有许多对阿基米德的描写.一些其他希腊和罗马历史学家也记述了阿基米德生平.

公元前 270 到 216 年,希罗 Ⅱ 统治着繁荣昌盛的叙拉古城,阿基米德的父亲是天文学家菲迪亚斯(Phidias),可能是叙拉古国王希罗 Ⅱ 的亲戚.阿基米德早年在亚历山大度过,在那里他发明了灌溉用的抽水机(图 3.2).他许多著作的前言都寄给了亚历山大的学者们,包括一位叫埃拉托塞尼(Eratosthenes)的图书馆馆长.阿基米德一生的大部分时间都是在故乡叙拉古度过的,在那里,希罗王和他的学生们反复要求他用他的数学才能来解决各种实际问题.很多故事记载了他对待工作的热情奉献精神.普卢塔克写道,在很多情况下,他对数学的专注"使他忘记了吃饭,并且忽视了个人的存在,甚至于达到了这样的程度,即当他被强行拉去洗澡或接受涂油礼(anointed)时,他仍完全沉浸在一种对科学神圣的热爱和乐趣中,经常在火堆的灰烬中或者在涂在他身体的油层上描绘图形."[4]

阿基米德军事工程师的天才使围攻叙拉古的马塞勒斯率领的罗马军队在海湾滞留数月之久.然而,最后可能是因为叛变,罗马军队攻进了叙拉古.马塞勒斯明令不准伤害阿基米德,但普卢塔克写道:"他正聚精会神地思考要解决的问题,以至根本没有觉察到罗马军队已经攻陷了这座城市.一个罗马士兵突然出现在他面前,命令他马上去见马塞勒斯,他拒绝了那个士兵的要求,表示要等解决了问题并给出证明,否则是不会去的.这个罗马士兵被激怒了,于是拔剑杀害了他."[5]

若按实际的情况,我们既要考虑放在两端点上的物体的重量和它们距支点的距离,还要考虑杠杆自身的重量和结构.杠杆自身两端不一定等重,其密度也许不同,若在杠杆上放了物体后,杠杆也许会轻微弯曲甚至折断;另外,支点也是一个有特定尺度的个体,杠杆或许会沿支点滑动,因此,不可能准确地确定出支点的位置及物体离支点的距离.在用数学方法分析杠杆时,上述这些因素使得对杠杆的数学研究变得十分困难.而阿基米德对物理状态进行简化,他假定杠杆是刚性的且自身无重量,支点和重物都是数学点,于是他可以给出杠杆的数学原理.



图 3.2 阿基米德.

在《论平面图形的平衡》(或《论平面图形的重心》) 的开始部分,阿基米德讲述了这些原理.受希腊几何的影响,他的这一著作以 7 个假定的公设为开端.这里列出其中的四个.

1. 距支点等距离的等重物体处于平衡状态,不等距离的等重物处于非平衡状态,并向距支点远的一方倾斜.

2. 平衡状态下,若在一个物体处加上一些重物,那么它们将变成非平衡状态,且向加重物的一方倾斜.

3. 同样,若从其中任一方取走一些重物,那么它们将失去平衡,且向未取走的一方倾斜.

6. 如果距支点一定距离的两个物体处于平衡状态,那么另两个与其同重的物体在距支点相同的距离处也处于平衡状态.

这些公设源自对杠杆的一些基本实验.实际上,第一个公设是我们通常所称的**"不证自明原理"**的一个例子.也就是说,我们假定等重物体在距支点等距平衡,这是因为我们没有理由作任何其它的假设.例如,杠杆不可能向右倾斜,因为从一个角度看是右方,从另一个角度来看则是左方.第二和第三个公设也是显然的.第六个公设乍看起来似乎没有意义.但是,用阿基米德的话说,第二句话的意思是指"其它的等重物体,若它们的重心位于与支点距离相同的两端,那么它们也处于平衡状态."也就是说,物体对杠杆平衡的影响只与它的重量和重心的位置有关.

虽然阿基米德著作中的标题及其它许多命题中都用了"重心"一词,但他从未给重心下一定义.很可能是他认为重心这一概念是众所周知的,没必要定义它.但在后来的希腊著作中却给出了重心的定义:"我们说,任意物体的重心是物体上的这样一点,如果把物体从该点悬挂起来,那么物体将静止不动,保持原来所处的位置."[6] 也许这个定义在阿基米德的时代已经被使用了.但阿基米德显然也已认识到 —— 正如他在公设 6 中所表达的 —— 物体向下的重力可以被认为是集中在物体的某一点上.阿基米德在其定理和命题中提及杠杆自身的重量时,把杠杆看作是没有重量的刚性物体,只考虑杠杆向哪一方倾斜.

由阿基米德所给出的 7 个公设中的前两个很容易推出杠杆原理:

命题 1 若距支点等距的两个物体处于平衡状态,则它们的重量相等.

命题 2 不等重的两物体在距支点等距处处于非平衡状态,且杠杆向重的一方倾斜.

第一个命题用归谬法证明.假设两物体重量不等,那么从较重的一方取走一些后,使它们变得相等.由公设 3,此时杠杆将失去平衡,这与公设 1 矛盾,故假设不成立.命题 2 的证明也是从较重一方取走一些后,使两边物体重量变得相等.由公设 1,此时杠杆处于平衡,如果再把取走的重量加到其中一方,那么杠杆将向较重的一方倾斜.



图 3.3 《论平面图形的平衡》命题 3.

命题 3 假设 A、B 两物体重量不等,且 $A > B$,它们相对于支点 C 保持平衡(图 3.3),令 $AC = a, BC = b$,则 $a < b$.相反

地，如果 $a < b$，则 $A > B$.

此命题亦可用归谬法证明．假设 $a \not< b$，从 A 中减去两边重物之差 $A - B$，由公设 3，杠杆将向 B 方倾斜．但如果 $a = b$，由于两边重量相等，此时杠杆处于平衡状态；如果 $a > b$，由公设 1，杠杆将向 A 方倾斜．这两种情况都与 $a \not< b$ 的情况矛盾，所以只有 $a < b$．同样可证相反的情况．

阿基米德在命题 4 和 5 中叙述了质量分布均匀的两个或三个等重物体构成的系统的重心在它们的几何中心上．在该命题中这一结论被推广到任意多个质量分布均匀的物体构成的系统，如果在离中心等距处重量相等．在命题 6 和 7 中则叙述了杠杆原理．

命题 6、7 可公度（命题 6）或不可公度（命题 7）的两个量处于平衡状态，当它们到支点的距离与量的大小成反比．

首先假设 A 和 B 是可公度的量，即 $A : B = r : s$，其中 r, s 皆为数．阿基米德断言，如果把 A 放在点 E，B 放在点 D，且在 DE 上取点 C，使得 $DC : CE = r : s$，那么点 C 便是 A 和 B 的重心（图 3.4）．要证这一结论，先选定单位使 $DC = r$，$CE = s$．然后在 DE 上取一点 H，使得 $HE = r$，延长 HE 到点 L，且 $EL = r$，再反向延长 HE 到点 K，且 $DK = HD = s$，那么点 C 是线段 LK 的中点．现在把物体 A 分成 $2r$、B 分成 $2s$ 等份，然后分别均匀地布放在 LH 和 HK 上．因为 $A : B = r : s = 2r : 2s$，故物体 A 的每一部分与物体 B 的每一部分相等，由前面的结论可知，物体 A 的重心在 HL 的中点 E 上，物体 B 的重心在 KH 的中点 D 上，由公设 6，这等同于把 A 放在点 E、B 放在点 D 上．另外，这个系统是由沿 KL 均匀分布的 $2r + 2s$ 个小部分组成，因此整个系统的重心在线段 KL 的中心点 C 上．故放于点 E 的物体 A 和放于点 D 的物体 B 相对于点 C 平衡．

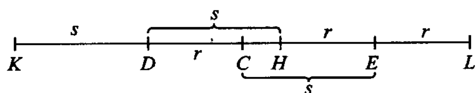


图 3.4 论平面图形的平衡命题 6.

对两个不可公度的量，阿基米德用约化（reductio）的方法得出相关结论．如果有两个不可公度的量，可以从其中一个物体中减去任意小的量后，使它变成与另一个量，可公度．有趣的是，阿基米德在这里既没应用《原本》卷 5 中不可公度量的欧多克斯比例理论，也没应用蒂奥泰德（Theaetetus）的以欧几里得算法为基础的该理论较早的形式．相反地，他主要是运用了一种连续性论证．虽然如此，他的证明仍不完善．

在这一著作的其余部分中，阿基米德用杠杆原理找到了不同几何图形的重心，证明了平行四边形的重心在其对角线的交点处，三角形的重心在其两条中线的交点处，抛物弓形的重心在其顶点到底线距离的 3/5 处．

3.1.2 工程应用

利用杠杆原理不仅可以得到几何结果，还可得到物理结果．特别地，对任意已知的两个物体 A、B 和一个杠杆，在杠杆上总存在一点 C，使 A、B 以 C 为支点保持平衡．如果 A 远重于 B，那么 A 在杠杆上充分接近 C，B 充分远离 C．若在 B 上任加一些物体，则 B 物体下降，A 物体上升．因此，阿基米德说了那句大话："我可以移动任何物体 …… 如果有另一个地球，我可以站在其上移动这个地球．"[7] 当希罗王听到这句话后，就请阿基米德以一个实际的实验演示他的原理．阿基米德并没用杠杆，取而代之的很可能是使用了某种滑轮或滑车系统，它们也可产生巨大的机械效益．普卢塔克（Plutarch）记述道："他从国王的军库中选定了一艘货船，如果没有大量劳工这只船是不可能拖出码头的；阿基米德让船载满旅客和货物，自己坐在远远的地方，手中握住滑轮的一端，然后轻松地拉动

绳索,轮船竟沿着直线缓缓前进,就像在海上行驶一样"[8].普卢塔克的故事还有一种略微不同的说法,大致是说阿基米德负责建造了一艘名为叙拉古的巨轮(the Syracusa),并用一只手就动了这个4200吨的宠然大物.

阿基米德也因设计了许多战争机械而非常出名.这些机械的使用使叙拉古抵抗罗马侵略长达好几个月.阿基米德还设计了许多火箭发射器和起重机,利用这些起重机可把罗马战船从海面上拖起再撞向岩石或把船员和战士抛出甲板.阿基米德的这些成功使罗马士兵胆颤心惊,他们一见到绳子或木片在城墙上出现便急忙逃跑.

普卢塔克认为,阿基米德不太乐于承认自己是一个工程师,"他不愿屈尊去对机械工程作评论或写这些方面的著作;他把整个工程技术看成是低级的行业,对一些急功近利的手艺不屑一顾,而是把全部的精力投入到了那些可能与世俗生活没有关系的纯粹的理论思考中."[9]尽管如此,资料表明,事实上阿基米德还是写了一些机械方面的著作,如《论球的制作》(On sphere Making)描述了他制造的天象仪(天体运行的机械模型),另外还有一本关于水钟的著作.

王冠与洗澡的偶然事件把阿基米德引入了一个全新的研究领域——流体静力学,他发现了流体静力学基本定理:如果把一个重于液体的固体放在液体中所称得重量要比它实际的重量小,这一重量差等于固体排开液体的重量.但是,当阿基米德注意到洗澡时排出浴池的水时,他又是如何想到身体的重量会减少呢?或许他觉得在水里时身体要比平时轻一些.

阿基米德研究杠杆的同时,对流体静力学进行了数学研究,他的著作《论浮体》(On Floating bodies)仍以简洁的公设开始,然后得出一系列结论,其中如任何液体的表面是一球面,其球心便是地球的中心.阿基米德用流体静力学的基本定理解决了王冠问题.希斯(Heath)据5世纪一首拉丁诗歌的描述,给出了阿基米德应用流体静力学基本定理解决王冠问题的一种方法.[10]假设重为 W 的王冠分别由重为 w_1 的黄金和 w_2 的白银组成,其中 w_1 和 w_2 未知,要测定王冠中金、银所占的比例.先确定王冠实际重量与把它放在水中所称得的重量差 F,这可以通过称被它排开的水的重量确定.再把重 W 的黄金放入水中称得的重量为 F_1,由此可知 w_1 的黄金排开的水的重量等于 $\frac{w_1}{W}F_1$.同样,重 W 的白银放入水中称得的重量为 F_2,由此可知 w_2 的白银排开的水的重量等于 $\frac{w_2}{W}F_2$.因此,$\frac{w_1}{W}F_1 + \frac{w_2}{W}F_2 = F$,这样,王冠中金银的比例即可由下式给出:

$$\frac{w_1}{w_2} = \frac{F - F_2}{F_1 - F}.$$

维特鲁维乌斯也提出了一个解决王冠问题的方法,他的方法与上述的略有不同,更直接地依赖于洗澡的故事,却没有应用流体静力学基本定理.他还记载说阿基米德发现金匠欺骗了国王.金匠的下场如何呢?不得而知.

3.2 阿基米德和数值计算

阿基米德的短篇著作《论圆的度量》与欧几里得著作的不同之处是它包括了一些数值计算.其中的第一个命题叙述了化圆为方问题的解答,即只要知道圆的周长,便可求出已知半径的圆的面积.

命题 1 一个圆的面积 A 等于一个两直角边分别是圆的半径和圆周长的直角三角形的面积.

阿基米德的这一结果与巴比伦人的等价，即 $A = (c/2)(d/2)$，其中 d 是圆的直径长度，c 是圆的周长．阿基米德用欧多克斯的穷竭法对此命题给出严格证明．先假设 $A > K$（K 是三角形的面积），在圆中内接一个边数不断增加的正多边形，最后得到一个面积等于 P 的正多边形，且满足 $A - P < A - K$，所以 $P > K$．而从圆心到多边形一边中点的垂线小于圆的半径，多边形的周长小于圆周长，故 $P < K$，与上面矛盾．同样，若假设 $A < K$，亦可得出矛盾．故命题成立．

命题 3 给出圆周长的一个近似数值来补充说明命题 1．

命题 3　圆的周长与直径的比小于 $3\frac{1}{7}$，而大于 $3\frac{10}{71}$．

阿基米德在命题 3 的证明过程中提出了计算圆内接正多边形和外切正多边形面积的一个算法，他先从正六边形入手，由初等几何的知识可知其边长与圆直径的比值．再应用以下引理依次计算边数为 12、24、48、96 的正多边形的周长与圆直径之比．

引理 1　设 OA 是圆的半径，CA 是过圆上点 A 的切线，$\angle COA$ 的平分线 DO 与 AC 相交于点 D，那么 $DA/OA = CA/(CO + OA)$，$DO^2 = OA^2 + DA^2$（图 3.5）．

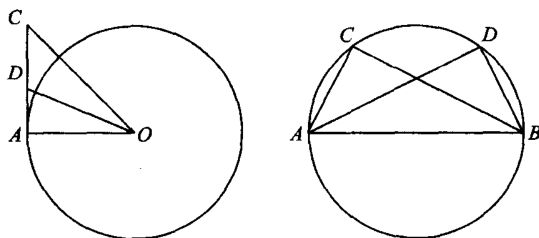


图 3.5　《圆的度量》引理 1 和 2.

引理 2　AB 是圆的直径，ACB 是内接于半圆的一个直角三角形，$\angle CAB$ 的角分线 AD 与圆周交于点 D，连接 DB，那么，$AB^2/BD^2 = 1 + (AB + AC)^2/BC^2$，$AD^2 = AB^2 - BD^2$．

为了得到要求的比，阿基米德考虑圆的外切正多边形的情形，重复应用引理 1 得出一个递推算法．先假设 $\angle COA$ 是直角的三分之一(30°)，那么 CA 等于圆外切正六边形边长的一半，故 CA 和 CO 已知．因为 $\angle DOA = 15°$，所以 DA 等于正十二边形边长的一半，由引理 1 可求出 DA 和 OA．再把 $\angle DOA$ 二等分得一个 $7\frac{1}{2}°$ 的角，这个角对应的线段长是正二十四边形边长的一半，而正二十四边形的边长也可算出．若 r 是圆的半径，t_i 是一个正 3×2^i 边形边长的一半，u_i 是圆心到该外切正多边形顶点的长，那么引理 1 可简化为以下递推式：

$$t_{i+1} = \frac{rt_i}{u_i + r}, \qquad u_{i+1} = \sqrt{r^2 + t_{i+1}^2}.$$

那么圆外切正 i 边形的周长与圆直径比为 $6(2^i t_i) : 2r = 3(2^i t_i) : r$．

对圆内接正多边形的情形，他也根据引理 2 得出了类似的算法．阿基米德计算出了这两种情况中每一步具体的数值，如在对内接和外切正六边形的计算中，须估计 $\sqrt{3} : 1$ 的近似值．他确定了这个近似值大于 265 : 163，而小于 1351 : 780．虽然我们不能确切地知道他是如何计算出这些值的，但有一点可以肯定的是，阿基米德像后来的数学家一样，非常擅长于数值计算．在上述两个算法中的第四步之后，他事实上已经得出了圆外切正 96 边形的周长与圆直径的比小于

$$14\,688 : 4673\frac{1}{2} = 3 + \frac{667\ 1/2}{4673\ 1/2} < 3\frac{1}{7},$$

而圆内接正 96 边形的相应比大于 $6336 : 2077\frac{1}{4} > 3\frac{10}{71}$,即得命题 3 的结果.

阿基米德的这一结果是历史上 π 值计算的最早记录.一旦有了确定的方法,只要有足够的耐心,我们便可以计算出 π 的任意精度的值.阿基米德没有说为什么他只计算到正 96 边形,而 $3\frac{1}{7}$ 这个值现在已成为 π 的一个标准近似值.

阿基米德的继承者尼古米德(Nicomedes,晚于公元前 3 世纪)采用了一种新的方法计算圆的周长,再应用命题 1 研究化圆为方问题.他用两直线的运动给出了上个世纪已引入的**割圆曲线**的定义:在正方形 $ABCD$ 中,令射线 AB 以 A 为定点从起始位置旋转到 AD,同时 BC 平行地移动到 AD 上(图 3.6),那么割圆曲线 BZK 是 AB 和 BC 同时移动过程中交点的轨迹.由此定义可知,割圆曲线上的点 Z 满足下面比例式:$ZL : BA$ = 弧 DG : 弧 BD,或 ZL : 弧 $DG = AB$: 弧 BD.用现在的术语可表示为:如果 $\rho = \rho(\theta)$ 是曲线的极坐标方程,那么 ρ 满足方程

$$\frac{\rho(\theta)\sin\theta}{a\theta} = \frac{a}{\frac{1}{2}\pi a},$$

其中 a 是正方形的边长.

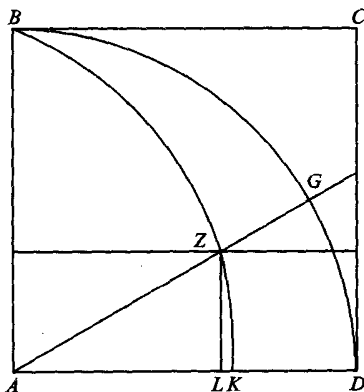


图 3.6　割圆曲线.

如果上式中取 θ 趋于 0 的极限,即可得到

$$\frac{\rho(0)}{a} = \frac{a}{\frac{1}{2}\pi a}.$$

事实上,古希腊人并未能给出上面的极限,但结果 $AK : AB = AB$: 弧 BD 很可能是被尼古米德用双归纳法证明的.由此可知,弧 BD(即四分之一圆周)是关于已知线段 AK 和 AB 的第三比例项,因此可以用欧几里得工具作出.(需指出的是,因为曲线的这一定义不能确定终点 K,所以即使在古代,这种作图是不被接受的.这只是一个近似解.)

3.3　阿基米德与几何

阿基米德的几何研究不同于欧几里得之处是在于他经常给出发现定理的方法或在严格的证明之前给出分析过程.在其专著《方法论》中收集了他的一些结果的发现方法.1899 年,这一著作的一个手抄稿被意外发现.该手稿的日期被确定为公元 10 世纪.但在 13 世纪时,稿中的一些字已被擦掉,而写上了与宗教相关的内容(在中世纪,羊皮纸是价值很高的商品).值得庆幸的是这些旧的字迹大部分仍然可以辨认.海伯哥(Heiberg)1906 年在君士坦丁堡(Constantinople)考察了这些原稿,不久后,他用希腊文发表了这些内容.

3.3.1　阿基米德的发现方法

《方法论》的内容包括阿基米德计算许多面积和体积的方法、技巧及一些重要结果,其中大部

分结果都有严格的证明.这些方法的基本特点是利用杠杆原理,把要计算面积的图形与已知图形对应地进行切割后,使它们在杠杆上保持平衡.因为这些机械的方法和"不可分的"截面都不能以正式的数学方式表达,所以阿基米德知道用这种方法并不能得出严格的证明.因此,他在《方法论》的序言中写道,这些定理"将来必须用几何的方法加以证明,因为用以上方法进行的研究并不算是真正意义上的证明".他还指出,"应用这种方法得出问题的结论后,再去寻求这些结论的证明当然比一无所知的情况下去发现这些问题结论容易得多."[11]

下面给出《方法论》中的命题1,把它作为阿基米德这一工作的一个典型.抛物弓形(的面积)是其内接三角形(面积)的4/3.阿基米德所说的抛物弓形是指由抛物线 ABC 和直线 AC 围成的图形,其中 B 称为抛物弓形的顶点,它是过 AC 的中点与抛物线轴平行的直线和抛物线的交点(图3.7).在抛物线上,点 B 到 AC 的垂直距离最大,点 B 称为**顶点**.已知抛物线 ABC 及其顶点 B,过点 C 作抛物线的一条切线与 BD 交于点 E,再过点 A 作平行于 DE 的直线与切线交于点 F,延长 CB 交 AF 于点 K,再延长到点 H,使得 $CH = KH$.阿基米德的思想是把 CH 作为以 H 为中点的杠杆,证明图中的 ΔCFA 与放在点 H 的抛物弓形保持平衡.过程如下,先在三角形 CHF 中任意作平行于 ED 的线段 MO,再证 MO 与放在点 H 的抛物弓形中的线段 PO 平衡.要证明二者平衡,须用到抛物线的两个性质,即 $EB = BD$ 和 $MO : PO = CA : AO$(显然阿基米德对抛物线的基本性质十分熟悉).因为 $EB = BD$,所以 $FK = KA$,$MN = NO$,又因为 K 是 AF 的中点,由《原本》卷6命题2可知,$MO : PO = CA : AO = CK : KN = HK : KN$.若把与 PO 相等的线段 TG 置于其中心 H 点,则有 $MO : TG = HK : KN$.由于点 N 是 MO 的重心,那么由杠杆原理可知 MO 与 TG 以 K 为支点保持平衡.

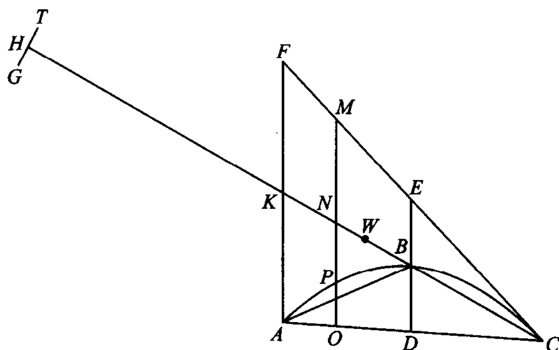


图3.7　平衡抛物弓形.

阿基米德接着写道,"因为 ΔCFA 由所有像 MO 这样的平行线段组成,而抛物弓形 CBA 由所有像 PO 一样含于曲线内部的线段组成,因而上图中的三角形与重心放在点 H 的抛物弓形关于支点 K 保持平衡."[12]若把三角形看作是一个重心点 W,这并不影响结论.由于 W 在 CK 上,且 $KW = \frac{1}{3}KC$,阿基米德推出 ΔACF : 抛物弓形 $ABC = HK : KW = 3 : 1$,即抛物弓形 $ABC = \frac{1}{3}\Delta ACF$,又因为 $\Delta ACF = 4\Delta ABC$,所以抛物弓形 $ABC = \frac{4}{3}\Delta ABC$.最后,阿基米德重点指出,"这里所陈述的事实不能以上面的过程作为真正的证明,但这一过程却暗示了结论的正确性.鉴于该定理并未得到证明,同时它的真实性又值得怀疑,因此我们将求助于几何学上的证明,我本人已经发现并公布了这一证明."[13]

有趣的是,阿基米德在这里使用了两个多世纪以前德谟克里特曾使用过的同样术语,即平面或

立体的不可分截线或截面组成了整个图形.虽然阿基米德在《方法论》中多次用到不可分量,但他没有说为什么应用它.这使我们相信,与阿基米德有联系和交往的亚历山大的学者,特别是数学家,都了解不可分量的应用,虽然他们也知道用这种方法研究问题并不能形成严格的几何证明,但他们在类似问题的研究中仍应用不可分量的方法.[14]

3.3.2 级数求和

阿基米德在其著作《求抛物弓形的面积》中的确给出了抛物弓形面积结果的几何证明,其证明方法即是欧多克斯的的穷竭法.阿基米德的想法是在弓形内作内接多边形,使弓形的面积与内接多边形的面积差小于任一给定值.先在弓形中作三角形 PQQ',点 P 把弓形分为 PRQ 和 $PR'Q'$ 两个部分,接着再作三角形 PRQ 和 $PR'Q'$,这两个三角形把弓形分为四个部分,依此类推地作下去(图 3.8).

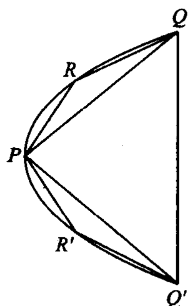


图 3.8 用几何级数求抛物弓形的面积.

阿基米德计算出每次作出的所有三角形的面积之和是前一次作出的所有三角形面积和的 1/4,所作次数越多,这些三角形的面积和越接近弓形的面积.为了完成这一证明,他需要成功地计算出几何级数

$$a+\frac{1}{4}a+\left(\frac{1}{4}\right)^2a+\cdots+\left(\frac{1}{4}\right)^na$$

的和,其中 a 为 $\Delta PQQ'$ 的面积.阿基米德并没有应用《原本》卷 9 命题 35 的几何级数求和公式,而是直接给出以下形式的结果:

$$a+\frac{1}{4}a+\left(\frac{1}{4}\right)^2a+\cdots\left(\frac{1}{4}\right)^na+\frac{1}{3}\left(\frac{1}{4}\right)^na=\frac{4}{3}a.$$

并用双归谬法(double reductio ad absurdum)给出了证明.假设 $K=\frac{4}{3}a$ 与弓形面积 B 不相等.如果 $K<B$,那么在弓形内可按上述方法作内接三角形,且满足 $B-T<B-K$,其中 T 是内接三角形的总面积.如果 $T>K$,由求和公式可得 $T<\frac{4}{3}a=K$,这是不可能的.另一方面,如果 $K>B$,可以确定 n 的值,使得 $\left(\frac{1}{4}\right)^na<K-B$.因为 $K-T=\frac{1}{3}\left(\frac{1}{4}\right)^na<\left(\frac{1}{4}\right)^na$,所以可知 $B<T$,这同样也是不可能的.综上可知,$K=B$.

这一证明的重要引理给出了几何级数求和的方法.阿基米德用一个含有五项的级数来表示这一结果.与欧几里得一样,阿基米德也没有把级数表示成任意多项的概念,但是我们很容易把他的方法推广.用现在的记号,n 表示任意正整数,阿基米德首先指出:

$$\left(\frac{1}{4}\right)^na+\frac{1}{3}\left(\frac{1}{4}\right)^na=\frac{1}{3}\left(\frac{1}{4}\right)^{n-1}a.$$

然后计算如下:

$$\begin{aligned}&a+\frac{1}{4}a+\left(\frac{1}{4}\right)^2a+\cdots+\left(\frac{1}{4}\right)^na+\frac{1}{3}\left[\frac{1}{4}a+\left(\frac{1}{4}\right)^2a+\cdots+\left(\frac{1}{4}\right)^na\right]\\&=a+\left(\frac{1}{4}a+\frac{1}{3}\cdot\frac{1}{4}a\right)+\left[\left(\frac{1}{4}\right)^2a+\frac{1}{3}\left(\frac{1}{4}\right)^2a\right]+\cdots+\left[\left(\frac{1}{4}\right)^na+\frac{1}{3}\left(\frac{1}{4}\right)^na\right]\\&=a+\frac{1}{3}a+\frac{1}{3}\cdot\frac{1}{4}a+\cdots+\frac{1}{3}\left(\frac{1}{4}\right)^{n-1}a\end{aligned}$$

$$= a + \frac{1}{3}a + \frac{1}{3}\left[\frac{1}{4}a + \cdots + \left(\frac{1}{4}\right)^{n-1}a\right].$$

化简整理后得要求的结果：

$$a + \frac{1}{4}a + \left(\frac{1}{4}\right)^2 a + \cdots + \left(\frac{1}{4}\right)^n a + \frac{1}{3}\left(\frac{1}{4}\right)^n a = \frac{4}{3}a.$$

在《论螺线》中，阿基米德也给出了一个面积计算的公式，同样也是用欧多克斯穷竭法予以证明. 该著作的命题 10 阐述了前 n 个正整数平方和的公式，

$$(n+1)n^2 + (1+2+\cdots+n) = 3(1^2+2^2+\cdots+n^2).$$

同时给出一个推论：

$$3(1^2+2^2+\cdots+(n-1)^2) < n^3 < 3(1^2+2^2+\cdots+n^2).$$

用这一推论，他计算了"阿基米德螺线"一转所围图形的面积. "阿基米德螺线"用现在的极坐标表示即为 $r = a\theta$. 在《论螺线》的命题 24 中，阿基米德证明了完整的一转螺线与终点半径 AL 围成的图形面积 R 等于以 AL 为半径的圆面积 C 的 1/3. 他指出，总可给面积为 R 的螺线图形作内接和外切图形，使它们的面积差小于任意给定的值 ε（图 3.9）（根据《原本》卷 10 命题 1），通过连续地进行二等分，我们总可以确定一个正整数 n，使半径为 AL、圆心角是 $\left(\frac{360}{n}\right)^\circ$ 的扇形的面积小于 ε. 然后在每个圆心角为 $\left(\frac{360}{n}\right)^\circ$ 的扇形与螺线所围区域中作内接和外切的圆弧，从而可得所有外切图形面积与所有内接图形的面积之差等于最初选定的扇形的面积，即小于 ε.

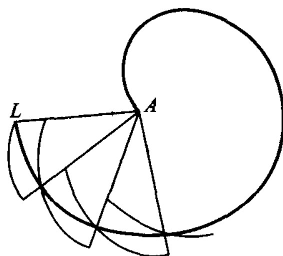


图 3.9　阿基米德螺线的面积.

下面给出这一结论的双归纳法证明（a double reductio）. 假设 $R \neq \frac{1}{3}C$，则有 $R < \frac{1}{3}C$ 或 $R > \frac{1}{3}C$. 对第一种情形，如上述那样作 R 的外切图形 F，使 $F - R < \frac{1}{3}C - R$. 因此，$F < \frac{1}{3}C$. 由螺线的方程可知，组成图形 F 的每个扇形的半径是一个算术级数，我们可以把它看作数列 $1, 2, \cdots, n$. 因为 $n \cdot n^2 < 3(1^2+2^2+\cdots+n^2)$，且这些扇形以及圆本身的面积与其半径的平方成正比，所以，$C < 3F$ 或 $\frac{1}{3}C < F$. 这与前面的 $F < \frac{1}{3}C$ 矛盾. 同样，可用作内接扇形的方法证明 $R > \frac{1}{3}C$，同样引出矛盾结论. 故命题得证.

3.3.3　分析法

最后我们要给出阿基米德工作的又一个例子，再次说明他不仅让读者知道几何问题的结论，还让他们知道得出这些结论的方法. 在《论球和圆柱》卷 2 命题 3 中，阿基米德阐述了作为证明背景的思考步骤.

问题　用一平面截给定的球，使得两球缺的面积比为已知比.

阿基米德对此问题的思考步骤或分析方法如下：先假定问题已解决，然后推导一系列结论，直至得到一个已知的真命题. 这样，他假设平面 BB' 将球分成两部分 BAB' 和 $BA'B'$ 且其面积比为 $H : K$（图 3.10）. 在《论球与圆柱》卷 1 中，他已证明了两球缺的面积分别等于以 AB、$A'B$ 为半径的圆的面积.

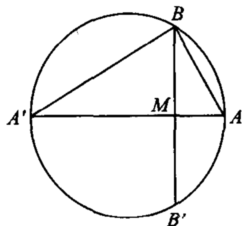


图 3.10　《论球和圆柱》命题 3.

所以,$AB^2 : A'B^2 = H : K$,且 $AM : A'M = H : K$(因两同高三角形面积之比等于底边长之比).而把一条线段按已知比分为两部分是已知的过程.阿基米德把上面这些步骤进行逆推,从而解决了原来提出的问题,即在 AA' 上取点 M,使 $AM : MA' = H : K$,那么由已知的定理可得 $AM : MA' = AB^2 : A'B^2 =$(以 AB 为半径的圆的面积):(以 $A'B$ 为半径的圆的面积)=(球缺 BAB' 的面积):(球缺 $BA'B'$ 的面积).从而问题可解.

阿基米德在卷 2 的命题 4 中给出了一个更复杂的分析过程.用一个平面截给定的球,使得两球冠体积之比为已知比.

阿基米德对此问题的分析如下:已知线段 ABC,且 $AB = 2BC$,点 E 在 BC 上,在 AB 上取一点 M,使得 $AB^2 : AM^2 = MC : EC$(图 3.11).若令 $AB = 2a, BC = a, EC = b$ 和 $AM = x$,此问题可以转化为一个代数方程:$(2a)^2 : x^2 = (3a - x) : b$ 或 $3ax^2 - x^3 = 4a^2b$.于是阿基米德需要解一个三次方程.他通过一条抛物线与一条双曲线的交点来确定所求的点 M,从而解决了问题.

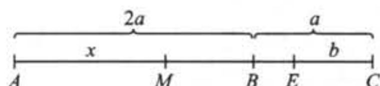


图 3.11 《论球和圆柱》命题 4.

以上只讨论了阿基米德现存的 14 部著作中的一小部分命题,这远不能充分说明他的数学天赋.阿基米德证明的其它结果包括:球的体积等于以球的大圆为底,半径为高的圆锥的体积的 4 倍;旋转抛物面的体积是与它同轴同底的圆锥体积的 3/2 倍;球的表面积是其大圆面积的 4 倍,等等.有充分的证据说明阿基米德还写了大量内容广泛的著作,涉及平衡、重心、半正多面体、光学、天文学等多个方面.罗马历史学家列维(Livy)把阿基米德说成是"一位无与伦比的天体和行星观测者."[15]



图 3.12 圣马力诺(San Marino)发行的阿基米德纪念邮票上印有一个以球的大圆为底的圆柱.

阿基米德去世后被安葬在叙拉古的一个城门附近.按他生前的要求,墓碑上刻着一个球及其外切圆柱的图形,以及他认为是他的最得意的定理,即以球的大圆为底,球的半径为高的圆柱,其体积是球体积的 3/2,其表面积也是球体表面积的 3/2(图 3.12).约公元前 75 年,西西里(Sicily)的执政官西塞罗(Cicero)发现了这座几近荒芜的坟墓,并进行了维修.虽然近几十年来不断有传闻说重新发现了阿基米德墓,但却没有充分的证据断定所发现的真是阿基米德墓.

3.4 阿波罗尼乌斯之前的圆锥曲线研究

圆锥曲线理论的起源甚为模糊,但它一定与倍立方体问题有关.如前所述,早在公元前 5 世纪,希波克拉底(Hippocrates)已经将倍立方体问题归结为求二次比的问题,即对一个棱长为 a 的立方体,在 a 和 $2a$ 之间确定 x 和 y,使得 $a : x = x : y = y : 2a$.用现在的术语来讲,这等同于同时解下列三个方程中的两个方程:$x^2 = ay$, $y^2 = 2ax$ 和 $xy = 2a^2$,前两个是抛物线方程,第三个是双曲线方程.

梅内赫莫斯(Menaechmus,公元前 4 世纪)是研究满足上述代数性质的曲线的第一人.他指出,通过这些曲线的交点可以确定 x 和 y,从而可以解决倍立方体问题.我们没有具体的资料确定他是如何研究这些曲线的,他很可能应用了欧几里得的方法.要确定曲线 $y^2 = 2ax$ 上的点,只能反复应

用《原本》卷 6 命题 13(图 3.13) 的方法. 先把长度为 $2a$ 和 x 的线段并为一个长度为 $2a+x$ 的线段,再以 $2a+x$ 为直径作一个圆,在两线段的结点处作一长为 y 的垂线(这时 y 满足方程). 如果取 x 的不同的值,并连接所有对应垂线的终点,即可作出所求的曲线[16]. 这里需指出,虽然可用欧几里得方法作出曲线上的每一点,但并不能作出完整的曲线. 不管怎样,作为一种工具,圆锥曲线理论已经被引入到解决某些几何问题中. 前面已谈到,阿基米德就曾用圆锥曲线解决有关球体的问题.

古希腊数学家怎样认识到他们用以解决倍立方体问题的曲线可以通过取圆锥的截线来产生呢?对此我们只能作些推测. 或许是梅内赫莫斯本人可能注意到图 3.13 中的圆图可被看作是某个圆锥的水平截线(level curves) 图,这可能使他想到这一特殊的曲线也可作为该圆锥的截线而得到. 另一种可能是这些曲线看上去就像太阳做周日圆周运动时晷影的运动路径,这路径本身是落在以日晷的头端为顶点的对顶圆锥的底面上. 按这种想法,投影所成的平面就是截平面. 进一步会注意到,从平面外一点看圆的外形像是一椭圆,它是用一个平面截视线圆锥所得的截面图形. 到公元前 4 世纪末,已有两本涉及圆锥曲线的著作,它们分别由阿里斯托斯(Aristaeus,公元前 4 世纪) 和欧几里得所著. 这两本著作均已失传,但从阿基米德关于圆锥曲线基本定理的广泛的参考文献中可以了解它们的内容.

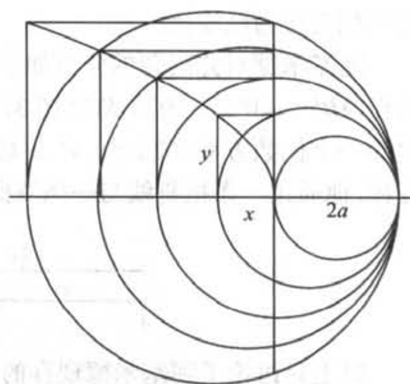


图 3.13　抛物线的欧几里得点作图.

欧几里得(《原本》卷 6) 把圆锥定义为以已知直角三角形一直角边为轴旋转所得的立体,然后按圆锥顶角大小把圆锥分为直角、钝角和锐角圆锥,用垂直于母线(已知直角三角形的斜边) 的平面截这个圆锥后得到一截线,"直角圆锥的截线(section of a right-angle cone)"、"锐角圆锥的截线(section of a acute-angle cone)" 和"钝角圆锥的截线(section of a obtuse-angle cone)" 即是现在所称的抛物线、椭圆和双曲线. 阿基米德及其后人已经使用了引号中的这些词.

3.5　阿波罗尼乌斯的圆锥曲线论

人物小传	阿波罗尼乌斯(250 B.C.—175 B.C.)
	阿波罗尼乌斯出生于小亚细亚南部的小城市波哥,有关他生平的记载资料很少. 我们所知道的他的生平情况基本来自其巨著《圆锥曲线》各卷(图 3.14) 序言中的记载. 从中可知他年轻时去亚历山大向欧几里得的学生求学,他一生的大部分很可能在那里讲学著书. 他的两种工作使他在古代非常出名,一是他的天文学工作,另一是他在后期的数学工作,其中大多数我们只是通过后人著作中的前言和摘要才略有所知. 幸运的是,他八卷本的《圆锥曲线》有七卷传了下来,这本著作在某种意义上说代表了古希腊数学发展的顶峰. 我们很难想像,在没有现代代数符号的情况下,他是如何发现并证明数百条优美深奥的定理. 毕竟,他完成了这些工作,其后的希腊数学著作没有一部在难度方面能超越《圆锥曲线》.

阿波罗尼乌斯在《圆锥曲线》中给出了与前人略有不同的三种圆锥曲线的定义. 他认为没必要限于用垂直于母线的平面截正圆锥来定义圆锥曲线. 他首先推广圆锥的概念如下:

APOLLONII
PERGAEI CONICORVM
LIBRI QVATTVOR.
VNA CVM PAPPI ALEXANDRINI
LEMMATIBVS, ET COMMENTARIIS
EVTOCII ASCALONITAE.
SERENI ANTINSENSIS
PHILOSOPHI LIBRI DVO
NVNC PRIMVM IN LVCEM EDITI.
QVAE OMNIA NVPER FEDERICVS
Commandinus Vrbinas mendis quamplurimis expurgata è Graeco conuertit, & commentariis illustrauit.

CVM PRIVILEGIO PII IIII. PONT. MAX.
IN ANNOS X.
BONONIAE,
EX OFFICINA ALEXANDRI BENATII.
M D LXVI.

图 3.14 阿波罗尼乌斯《圆锥曲线》拉丁文第一版的扉页，1566. 资料来源：Smithsonian 研究所图书馆，图号：86－4346.

> 从与圆不在同一平面上的一点作与圆周相交的直线，如果该点固定，把所作直线沿圆周旋转，……，那么生成的曲面是**一圆锥面**，固定点是**顶点**，顶点到圆心的直线是**轴**，圆称作圆锥的**底**.[17]

阿波罗尼乌斯所作的圆锥面即现在的对顶斜圆锥. 一般地，对顶斜圆锥的轴线与底圆并不垂直. 为简单起见，下面我们只讨论轴线与底面垂直的情况.

阿波罗尼乌斯先用过轴线的平面截圆锥定义三种曲线. 当此平面与底圆的交线 BC 是底圆的直径时，那么称截得的三角形 ABC 为轴线三角形. 与底圆相交于 ST，且 ST 垂直 BC 或 BC 的延长线的平面截圆锥后所得的曲线分别定义了抛物线、椭圆和双曲线(图 3.15, 3.16 和 3.17). 令 EG 是平面与轴线三角形的交线，如果 EG 平行于轴线三角形的一边，那么所得截线是抛物线；若 EG 与轴线三角形的两边相交，则截线是椭圆；最后，若 EG 与轴线三角形的一边以及另一边 CA 的延长线相交，则截线是双曲线，此时所得的曲线有两个分支，这与以往钝角圆锥的情形有所不同.

对上述各种情况，阿波罗尼乌斯分别导出了它们的"标征"，即曲线上任意一点的横、纵坐标的关系，由此可以容易地把曲线表示成一代数方程. 他在曲线上任取一点 L，过点 L 作平行于圆锥底面的平面，记所得截圆的直径为 PR，它与线段 EG 交于点 M，那么 LM 垂直于 PR，并且 $LM^2 = PM \cdot MR$.

如果 EG 平行于轴三角形的边 AC，阿波罗尼乌斯推导了抛物线的性质，即曲线上任意一点 L 的横、纵坐标 EM 和 LM 之间的关系(卷 1 命题 11). 为此，他作 EH 垂直于 EK(图 3.15)，使得

$$\frac{EH}{EA} = \frac{BC^2}{BA \cdot AC}.$$

图 3.15 抛物线方程的得出.

上面方程右端可写成$\frac{BC}{BA}$与$\frac{BC}{AC}$之积,但相似的,

$$\frac{BC}{BA}=\frac{PR}{PA}=\frac{PM}{EP}=\frac{MR}{EA},\quad \frac{BC}{AC}=\frac{PR}{AR}=\frac{PM}{EM},$$

所以,

$$\frac{EH}{EA}=\frac{ER\cdot PM}{EA\cdot EM}.$$

同时,

$$\frac{EH}{EA}=\frac{EH\cdot EM}{EA\cdot EM}.$$

因此,$MR\cdot PM=EH\cdot EM$且$LM^2=EH\cdot EM$.若设$LM=y,EM=x,EH=p$,则可得出抛物线的标准方程:$y^2=px$."parabola"一词源自希腊语 paraboli(贴合),这是因为以纵坐标y为边的正方形与贴合于横坐标x上的矩形相等.常数p称为抛物线的**参数**,它只与确定圆锥曲线的截面有关.

对其它两种情况,设D是EG与轴线三角形的另一边(椭圆情形)或其延长线(双曲线情形)的交点(图3.16和3.17).阿波罗尼乌斯对这两种情况证明了:以LM为边的正方形的面积,等于在线段EH上贴合的、宽为EM的矩形的面积,但有盈余(yperboli)或亏缺(ellipsis),盈、亏量等于一与以DE和EH为边的矩形相似的矩形面积,这也说明了这两种曲线名称的由来.他先在EH上作垂线ED,使得

$$\frac{DE}{EH}=\frac{AK^2}{BK\cdot KC}.$$

图3.16 椭圆标征的推导.

如上所述,方程的右边可写成积$\frac{AK}{BK}\cdot\frac{AK}{KC}$,这里$AK$与$DE$平行.类似地,

$$\frac{AK}{BK}=\frac{EG}{BG}=\frac{EM}{MP},\quad \frac{AK}{KC}=\frac{DG}{GC}=\frac{DM}{MR}.$$

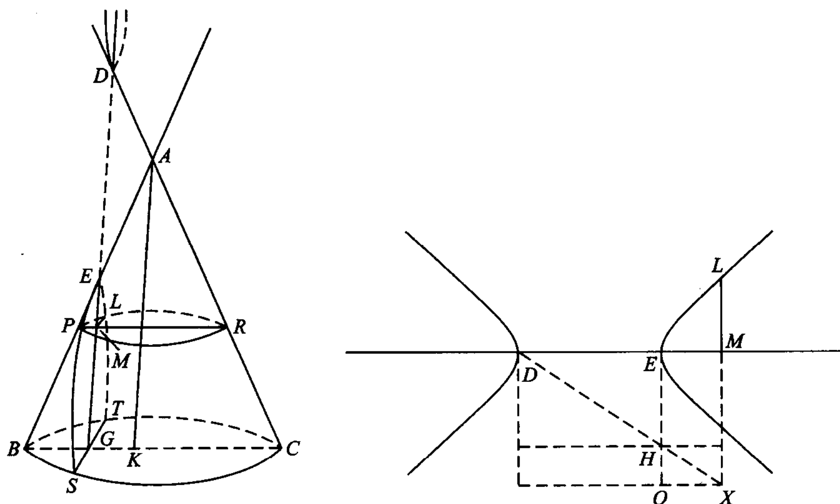


图 3.17 双曲线标征的推导.

因此，$\frac{DE}{EH}=\frac{EM\cdot DM}{MP\cdot MR}$. 但同时有$\frac{DE}{EH}=\frac{DM}{MX}=\frac{DM}{EO}=\frac{EM\cdot DM}{EM\cdot EO}$. 由此可得

$$MP\cdot MR=EM\cdot EO,\quad \text{故}\quad LM^2=EM\cdot EO.$$

在双曲线情形，$EO=EH+HO$，而在椭圆情形，则 $EO=EH-HO$. 这两种情况下，因为以 EM（$=OX$）和 HO 为长和宽的矩形与以 DE 和 EH 为长和宽的矩形相似，阿基米德证明了他的结论. 用现代的记法表示，因 $EM/HO=DE/EH$，故 $HO=EM\cdot EH/DE$，所以，若令 $LM=y$，$EM=x$，$EH=p$，$DE=2a$，那么阿波罗尼乌斯所得的标征就成为现代的方程：

$$y^2=x\left(p+\frac{p}{2a}x\right),\quad y^2=x\left(p-\frac{p}{2a}x\right),$$

它们分别是双曲线和椭圆的方程，其中参数 p 只与截面有关.

在得到椭圆和双曲线上面形式的方程后，阿波罗尼乌斯在卷 1 命题 12 中证明了两种曲线都可以写成 $y^2=\frac{p}{2a}x_1x_2$ 的形式，这里 x_1、x_2 分别是曲线的轴线上 ME 和 MD 的长度. 对椭圆而言，如果点(x,y)是短轴（长为 $2b$）的端点，则该方程表明有 $b^2=\frac{p}{2a}a^2$ 或 $b^2=\frac{pa}{2}$. 这样即可得出椭圆方程中参数 p 与其长轴和短轴长度间的基本关系. 下面将看到对双曲线此方程同样适用，这时 b 是双曲线顶点到渐近线的垂直距离.

通过圆锥曲线的定义得出这些标征后，阿波罗尼乌斯在卷 1 的最后证明了相反的命题：已知顶点（或一对顶点）、一个参数，可以确定一个圆锥和一个截面，而此截面在圆锥上所截得的曲线将是具有已知顶点、轴和参数的抛物线（椭圆或双曲线）. 从此以后，在希腊几何以及中世纪和早期近代几何学中，已知顶点、轴和参数，求作一条圆锥曲线就成了典型的问题，如同已知圆心和半径作圆一样. 这样，欧几里得《原本》增加了一些新的基本作图.

阿波罗尼乌斯在研究圆锥曲线的性质时，类似于现代的研究. 它通常不用曲线的原始定义，而是用它们的标征. 虽然阿波罗尼乌斯尽量使用几何的语言，但是其著作中有许多内容具有几何代数

的特点.也就是说,曲线标征的应用可以看作他几何研究的代数特点.因此,简单地说,《圆锥曲线》的杰出之处在于应用代数方法简化了几何问题的陈述和证明.

阿波罗尼乌斯在卷 2 中研究了双曲线的渐近线问题.在命题 1 中给出了渐近线的作法(图 3.18).在双曲线一顶点 A 作切线,在切线上取关于 A 对称的两点 L, L',使得

$$AL^2 = AL'^2 = \frac{pa}{2} \quad (= b^2),$$

过双曲线中心点 C 作直线 CL 和 CL',他证明了 CL 和 CL' 与曲线的两个分支不相交(希腊词 asymptotos 意思是"不相遇").命题 14 进一步说明了,如果这两直线无限延长,那么渐近线与双曲线之间的距离将小于任意给定值.

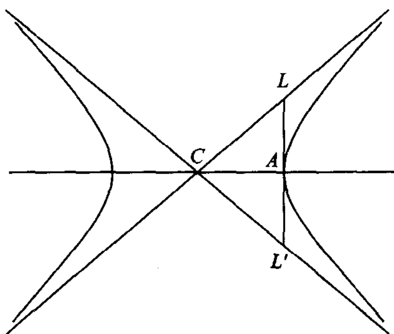


图 3.18 双曲线的渐近线作法.

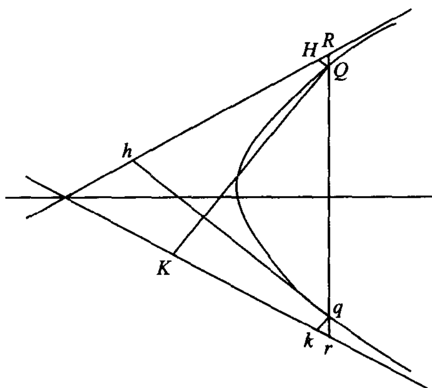


图 3.19 由渐近线定义的双曲线.

同时卷2命题4给出了已知双曲线上一点及其渐近线时作出该双曲线的方法.这提供了又一个基本作图.命题8指出双曲线的正割线在双曲线与其渐近线之间截得的线段长度相等.命题10和命题12叙述了双曲线的标征可不用参数和轴表示,而用其渐近线来表示.因为 $AL = b, AC = a$,我们以点 A 为坐标原点,那么渐近线方程可写为

$$y = \pm \frac{b}{a}(x + a).$$

再在双曲线上取两点 Q 和 q,使得 Qq 垂直于轴(图 3.19).如果 R 和 r 分别是直线 Qq 与两条渐近线的交点,记 $Q = (x, y)$,那么,由于 $b^2 = \frac{pa}{2}$,所以

$$\begin{aligned} QR \cdot Qr &= \left(\frac{b}{a}(x + a) - y\right)\left(\frac{b}{a}(x + a) + y\right) \\ &= \frac{b^2}{a^2}(x + a)^2 - y^2 \\ &= \frac{b^2x^2}{a^2} + \frac{2b^2ax}{a^2} + b^2 - px - \frac{p}{2a}x^2 \\ &= \left(\frac{b^2}{a^2} - \frac{p}{2a}\right)x^2 + \left(\frac{2b^2}{a} - p\right)x + b^2 = b^2. \end{aligned}$$

同理,可得 $qr \cdot qR = b^2$.如果过点 Q 、q 分别向两条渐近线各作一对平行线,它们与两条渐近线的

交点分别是 H、h 和 K、k，从图 3.19 可以看出，$RQ : Rq = HQ : hq$ 且 $qr : Qr = qk : QK$，由此即可得 $HQ : hq = qk : QK$ 或 $HQ \cdot QK = hq \cdot qk$，也就是说，双曲线的任一给定分支上的点到两条渐近线的距离的乘积是一常数. 这一结果用现在的符号表示，即双曲线可由方程 $x \cdot y = k$ 定义.

3.5.1 切线和法线

《圆锥曲线》卷 1 中讨论了作圆锥曲线切线的问题.

命题 33 已知抛物线 CET 上一点 C，且 CD 与抛物线的轴 EB 垂直，延长 BE 至 A，且 $AE = ED$，那么 AC 与抛物线相切于点 C(图 3.20).

设 $DC = y, DE = x, AE = t$，这一命题即可表述为：如果 $t = x$，那么 AC 是抛物线的过点 C 的切线，或者说，把抛物线的轴 BE 延长至点 A，且 $AE = x$，连线 AC 即是抛物线的切线. 阿波罗尼乌斯同欧几里得一样，把切线看作是与曲线相交但不切割曲线的直线. 他用反证法证明了命题 33. 假设直线 AC 与曲线上的另一点 K 也相交，那么线段 CK 在双曲线内部，在 CK 上取一点 F，过点 F 作轴线的垂线 FB，延长 BF 交曲线于点 G，则得 $BG^2 : CD^2 > BF^2 : CD^2 = AB^2 : AD^2$，又因为 G 和 C 同为曲线上的点，故 $BG^2 = p \cdot EB$ 且 $CD^2 = p \cdot ED$，于是 $BG^2 : CD^2 = BE : DE$，那么 $BE : DE > AB^2 : AD^2$，同时 $4BE \cdot EA : 4DE \cdot EA > AB^2 : AD^2$，即

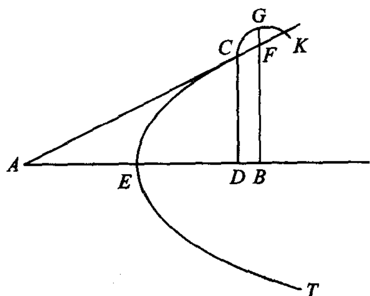


图 3.20 《圆锥曲线》卷 1 命题 33.

$$4BE \cdot EA : AB^2 > 4DE \cdot EA : AD^2.$$

《原本》卷 2 命题 5 指出，对任意长度 a、b，总有 $ab \leqslant \left(\frac{a+b}{2}\right)^2$ 或 $4ab \leqslant (a+b)^2$，当且仅当 $a = b$ 时等号成立. 所以由 $AE = DE$ 得 $4DE \cdot EA = AD^2$，由 $AE < BE$ 得 $4BE \cdot EA < AB^2$，所以在前面的不等式中，左边的值小于 1，而右边的值等于 1，得出矛盾.

命题 34 中给出了椭圆或双曲线的切线的作法，其证明过程与上面抛物线的证法相似.

命题 34 C 为椭圆或双曲线上的一点，CB 是点 C 到轴线上的垂线，设 G、H 是轴线与曲线的两个交点，在 GH 或其延长线上取一点 A，使得 $AH : AG = BH : BG$，那么 AC 即是过曲线上点 C 的切线.

这一结果可用如下代数语言描述：

设 $AG = t, BG = x$，对椭圆而言，$BH = 2a - x$、$AH = 2a + t$，对双曲线而言，$BH = 2a + x$、$AH = 2a - t$(图 3.21). 于是对椭圆有 $(2a + t)/t = (2a - x)/x$，对双曲线有 $(2a - t)/t = (2a + x)/x$. 由此可得 $t = ax/(a - x)$(对椭圆)或 $t = ax/(a + x)$(对双曲线)，这样便可作出它们的切线. 阿波

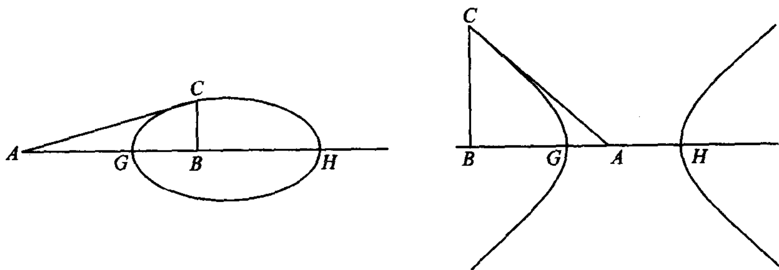


图 3.21 《圆锥曲线》卷 1 命题 34.

罗尼乌斯在命题35和36中证明了上面这两个命题的逆命题也成立,从而完整地解决了圆锥曲线的切线问题.

由于椭圆和双曲线的方程本质上是相同的,所以阿波罗尼乌斯经常对这两种曲线的性质采用相似的证明,甚至把它们放在同一个命题中叙述,对抛物线的性质却是单独讨论.

《圆锥曲线》第5卷中,阿波罗尼乌斯讨论了圆锥曲线的法线问题.以下只叙述抛物线的情形.

命题8、13、27 已知以点A为顶点的抛物线的方程是$y^2 = px$,G是轴上一点且$AG > \frac{p}{2}$,在AG上取点N,且$NG = \frac{p}{2}$.如果过点N作轴线的垂线NP与抛物线交于点P,那么PG是点G到曲线上的最短线段;相反地,如果PG是点G到曲线上的最短线段,那么PN垂直于轴线,且$NG = \frac{p}{2}$,同时PG与过曲线上点P的切线垂直(图3.22)[18].

设P'是以AN'为横坐标的抛物线上的另一点,由抛物线的定义可知,$P'N'^2 = p \cdot AN' = 2NG \cdot AN'$,又因为$N'G^2 = NN'^2 + NG^2 \pm 2NG \cdot NN'$(正负号与$N'$的位置有关),把这两个等式相加再由毕达哥拉斯定理可得$P'G^2 = 2NG \cdot AN + NN'^2 + NG^2 = PN^2 + NG^2 + NN'^2 = PG^2 + NN'^2$.故$PG$是点$G$到曲线上的最短线段.逆命题可用反证法得证.最后再证PG与切线TP垂直:因为$AT = AN$,那么$NG : p = \frac{1}{2} = AN : NT$,所以$TN \cdot NG = p \cdot AN = PN^2$,因为$PN$垂直$TG$,所以$\angle TPG$是一直角.

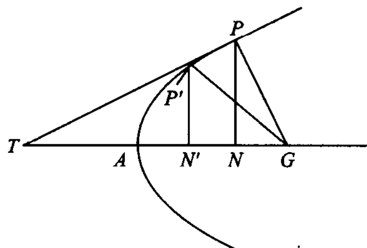


图3.22 《圆锥曲线》卷5命题8、13、27.

3.5.2 焦 点

下面给出《圆锥曲线》卷3中关于椭圆和双曲线焦点的一些结论.例如在卷3命题45中,阿波罗尼乌斯给出了椭圆焦点的定义:它们是椭圆长轴AB上的点F、G,满足AF和BF构成的矩形面积等于AB和参数p构成的矩形面积的$\frac{1}{4}$,且同样的结论对矩形AG和GB也成立(阿波罗尼乌斯称这些点为由轴线上的"矩形贴合所引出的点",1604年开普勒首先使用"焦点"一词).如果用代数语言表述,若设点F和G到椭圆中心O的距离都等于c,阿波罗尼乌斯的条件可写成下面的方程:

$$(a-c)(a+c) = \frac{1}{4} \cdot 2ap \quad \text{或} \quad a^2 - c^2 = \frac{ap}{2}.$$

在焦点定义的基础上,阿波罗尼乌斯给出了一系列命题,其中最为漂亮的结论是两焦点与椭圆上任一点的两条连线与椭圆在该点处的切线所成的角相等.

阿波罗尼乌斯也给出了双曲线类似的性质,但却没有给出抛物线焦点的性质,可能是在他的现已佚失的早期著作中讨论过这些性质.与阿波罗尼乌斯同时代的狄俄克利斯(Diocles,公元前2世纪早期)的著作《论火镜(On Burning Mirrors)》或许给出了抛物线焦点的以下性质:抛物线上任一点的切线与该点和焦点的连线所成的角等于此切线与平行于轴线的直线所成的角[19].狄俄克利斯的这一著作可能比《圆锥曲线》早若干年,他也可能是给出抛物线这一性质证明的第一人.此著作的名字正是源自抛物线的性质.怎样找到一个镜面,它可以把太阳光反射后集中于一点并使该点燃烧.狄俄克利斯证明了用旋转抛物线得到的抛物面可以实现这一点.曾有阿基米德和其他人用这样的镜子使敌船着火的故事,但我们毕竟没有可靠的证据证实这些故事的真实性.

为了全面讨论焦点问题,我们这里介绍狄俄克利斯著作中命题1对抛物线焦点性质的证明.已

知以 BW 为轴的抛物线 LBM,在 BW 上作线段 BE,且 BE 的长度等于抛物线参数 p,取点 D 为 BE 的中点(图 3.23),那么 D 即是抛物线的焦点(focus),它到顶点 B 的距离是 $p/4$.在抛物线上任取一点 K,过点 K 作切线 AC,AC 与 WB 的延长线交于点 A,再过 K 作平行于轴线的直线 KS,连接 DK.那么 $\angle AKD = \angle SKC$.

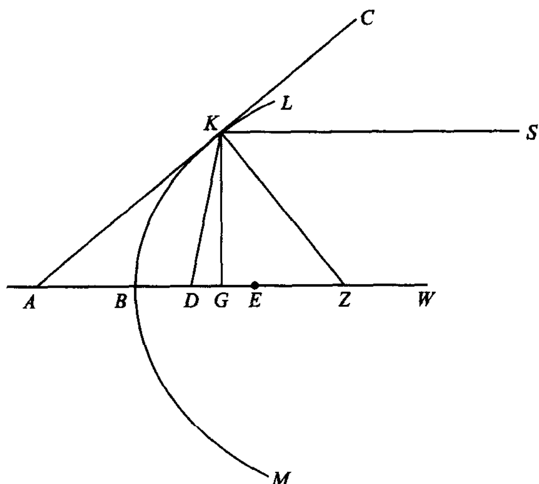


图 3.23　狄俄克利斯的《论火镜》.

要证明这一结论,先过点 K 作一条与轴垂直的直线且与轴线相交于 G,由《圆锥曲线》卷 1 命题 33 可知,$AB = BG$;再过点 K 作垂直于 AK 的直线且与轴线的交点为 Z,因为 $KG^2 = AG \cdot GZ$,同时 $KG^2 = p \cdot BG$,所以 $GZ = p/2$.又因为 $GZ = BE$,故 $GB = EZ$,$AB = EZ$,最后可得 $AD = DZ$.既然 D 是直角三角形 AKZ 斜边的中点,那么 $AD = DK = DZ$,因此 $\angle DZK = \angle DKZ$,又 $KS \parallel AZ$,则 $\angle ZKS = \angle DKZ$.再从两直角 ZKC 和 ZKA 中分别减去上面的两个等角后,即可得出要证的结论.狄俄克利斯最后指出:以 AZ 为轴旋转曲线 LMB 后得一曲面,把曲面的内侧镀铜后便可制成所谓的"火镜".

狄俄克利斯在命题4和5中叙述了已知焦距作抛物线的方法,此作法实际上应用了抛物线焦点和准线的性质.**准线**是这样的一条直线,抛物线上的任意一点到该直线的距离等于该点到焦点的距离.虽然公元 4 世纪的评论家帕普斯对此曾作过讨论,但迄今尚未发现有更早的论述抛物线的这一性质的文献.帕普斯还曾指出,到定点(焦点)与定直线(准线)距离之比是小于1的常数的点的运动轨迹是一椭圆;若这个常数大于 1,则是双曲线.大约在狄俄克利斯和阿波罗尼乌斯时期,人们可能已经发现了这些性质.

《圆锥曲线》中只讨论了椭圆和双曲线焦点的性质.卷 3 命题 51 叙述道,如果把双曲线上任意一点与两焦点相连,那么"长线段与短线段的差正好等于轴的长度";命题 52 则叙述道,在椭圆情形,这两条线段之和等于轴的长度.也就是说,若 P 是曲线上一点,D、E 是曲线的两个焦点,则对双曲线而言,$PD - PE = 2a$,对椭圆而言,$PD + PE = 2a$.这个性质正是现在的教材中这两种曲线的标准定义.

3.5.3　圆锥曲线在解决问题中的应用

在《圆锥曲线》中,阿波罗尼乌斯的目的并不纯粹在于研究圆锥曲线内在的优美性质,而是为了应用这些曲线来得到解决几何问题所必要的定理.因此,下面我们给出古希腊圆锥曲线应用的三

个例子作为本章的结尾.

先考虑三等分角问题.假设要三等分 $\angle ABC$(图3.24),过 A 作 BC 上的垂线,再作矩形 $ADBC$,延长 DA 到点 E,如果 BE 与 AC 交于点 F,那么线段 $FE = 2AB$.由此可得 $\angle FBC = \frac{1}{3}\angle ABC$.因为如果 G 为 FE 的中点,则 $FG = GE = AG = AB$,所以 $\angle ABG = \angle AGB = 2\angle AEC = 2\angle FBC$,即 $\angle ABC$ 被三等分.但是,要证明这一结论,必须说明怎样作出满足上述条件的线段 BE.这里借助了分析法,设 $FE = 2AB$,作 CH 和 EH,使得 $CH \parallel FE$, $EH \parallel AC$,那么点 F 在以 C 为圆心、$FE(=2AB)$ 为半径的圆周上,又因为 $DE : DB = BC : CF$,或 $DE : AC = DA : EH$,我们有 $DA \cdot AC = DE \cdot EH$,所以点 H 也在以 DB 和 DE 为渐近线且过点 C 的双曲线上.因此,如果作出上述的双曲线和圆,且过它们的交点 H 作一条垂直于 DA 延长线的线段 HE,那么垂足 E 即是解决三等分角问题所需要的点.

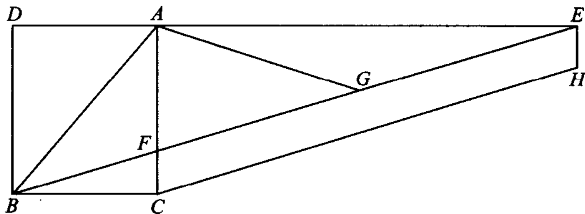


图3.24 用圆锥曲线三等分角.

倍立方体的作图本质上发始于希波哥拉底将该问题归结为求作已知线段 AB 和 AC 的二次比例中项的问题.有一个出现在阿波罗尼乌斯时代的倍立方体作图就是从以上述两线段为边作矩形着手的.作对角线 AD,以 AD 为半径作过点 B 的圆(图3.25),设 F 是圆和以 AB、AC 为渐近线且过点 D 的双曲线的交点,DF 的延长线分别和 AB、AC 的延长线交于 E、G 点.由《圆锥曲线》卷2命题8可知,$FE = DG$,故 $DE = FG$.又因为 F、D、C、A、B 是同一个圆上的点,由《原本》卷3命题36可知,$GA \cdot GC = GF \cdot GD$ 和 $EA \cdot EB = ED \cdot EF$,因此,$GA \cdot GC = EA \cdot EB$ 或 $GA : EA = EB : GC$.类似地,$GA : EA = DB : BE = AC : BE$,同时 $GA : EA = GC : DC = GC : AB$,所以 $AC : BE = BE : GC = GC : AB$,即 BE、GC 是所求的两个比例中项.

最后一个问题是三、四线轨迹问题(three-and four-line problem),直到17世纪,这一问题仍是数学家们热烈讨论的话题.问题的基本形式可陈述如下:

已知三条直线,确定一点,使得该点到其中一条直线的距离平方与到另两条直线距离的乘积之比为一常数(这里的每一条直线的距离是分别按给定的角度来测定的).这里考察一个特例,如果其中两条直线平行,第三条直线与前两条直线垂直,那么很容易分析地看出满足条件的点的运动轨迹是圆锥曲线.前面讲到,椭圆和双曲线的方程可以写成 $y^2 = \frac{p}{2a}x_1x_2$,这里 y 是曲线上的点到轴线的距

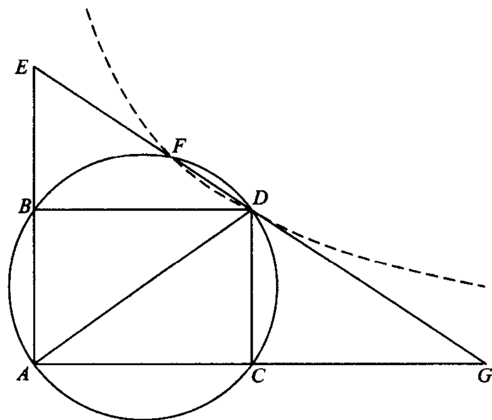


图3.25 用圆锥曲线解倍立方体问题.

离，x_1、x_2 是该点到曲线轴的两个端点的横坐标距离. 如果在这两个端点作曲线的切线，那么此圆锥曲线就是由轴线和两条切线构成的三线轨迹问题的解.

而希腊数学家们想的是怎样把这一结果推广，即证明无论三条直线之间相互位置如何，动点的轨迹总是圆锥曲线. 阿波罗尼乌斯曾叙述道（见本章开始的引语），欧几里得只是部分地解决了三线轨迹问题，而他自己在《圆锥曲线》卷 3 中的新结论可以完整地解决三线轨迹问题. 虽然阿波罗尼乌斯在卷 3 中并没有以上述形式提出这一问题，事实上从其中的一些定理我们可以知道，从圆锥曲线外一已知点所作的曲线的两条切线与连接两切点的直线（曲线的割线）构成一个三线问题，而此圆锥曲线就是该问题的解. 通过著作中的另外一些定理可以看出圆锥曲线也可用来解决四线轨迹问题，即确定到两条直线的距离之积与到另两条直线的距离之积的比值是一常数的点的轨迹. 在古希腊晚期，一些数学家试图解决更多直线的轨迹问题，但均未有大的进展. 直到 17 世纪，笛卡儿（Descartes）和费马（Fermat）仔细阅读阿波罗尼乌斯的著作后，用解析几何的方法解决了这一问题. 笛卡儿能够对任意数目的直线推导出满足类似条件的曲线的方程，并且对解进行分类. 我们运用现代符号对许多希腊问题进行的描述分析清楚地表明了：希腊人解决几何问题的传统，在希腊世界经过长久的沉寂之后，被阿拉伯世界继承，并最终导致了数学方法的新的进步，而这种进步使得大多数这类希腊问题被简化成为教科书里的习题.

习　题

阿基米德对平面平衡的研究

1. 在长为 10 米的杠杆上找一支点，使得两端分别放 14 千克和 10 千克的物体后杠杆平衡.
2. 在距杠杆支点 10 米处放一 8 千克的物体，距支点 8 米的另一端放 12 千克的物体后，杠杆将向哪方倾斜？
3. 证明命题 6 的逆命题.

命题 6 的逆命题　如果两重物在杠杆上平衡，那么它们到杠杆支点的距离与其重量成反比.

阿基米德对浮体的研究

4. 维特鲁维乌斯在《建筑学》（On Architecture）中给出了阿基米德解决王冠问题的另一种可能的方法，假设重为 W 王冠分别由重为 w_1 和 w_2 的黄金的白银组成，把王冠投入液体中排开液体的体积是 V，与王冠等重的黄金和白银投入液体中分别排开液体的体积是 V_1 和 V_2，求证

$$V = \frac{w_1}{W}V_1 + \frac{w_2}{W}V_2, \quad \text{因此} \quad \frac{w_1}{w_2} = \frac{V_2 - V}{V - V_1}.$$

阿基米德对圆的度量的研究

5. 证明阿基米德计算 π 的算法中所用的两个引理.
6. 利用计算器（或计算机）重复应用阿基米德引理 1 中的算法计算 π 的值. 若要精确到五位小数，则需重复迭代多少次？
7. 把引理 2 写成递推算法，重复应用这一算法计算 π 的精确到五位小数的值，则需迭代多少次？
8. 证明：如果 a 是与 $\sqrt{a^2 \pm b}$ 最接近的正整数，则

$$a \pm \frac{b}{2a} > \sqrt{a^2 \pm b} > a \pm \frac{b}{2a \pm 1}.$$

由 $2^2 - 1 = 3$ 开始,试证:第一个近似值 $2 - 1/4 > \sqrt{3} > 5/3$;第二个,$\sqrt{3} < 26/15$;第三个,$\sqrt{3} < 1351/780$;第四个,$\sqrt{3} > 265/153$.后面的两个近似是阿基米德在《圆的度量》中使用的值.

阿基米德的《方法论》

9. 已知 MO 与抛物弓形的轴平行,MC 与抛物线相切(见图3.7),试分析证明:$MO : OP = CA : AO$.

10. 阿基米德《方法论》中命题2的证明提纲:

命题2　球(就体积而言)是以它的大圆为底、它的半径为高的圆锥体积的4倍.

设 $ABCD$ 为球的大圆,AC、BD 是相互垂直的直径.作一个以 A 为顶点、AC 为轴的圆锥,扩展这一圆锥的表面到以 EF 为直径的圆.以该圆为底,AC 为高和轴作一圆柱,并延长 CA 至 H,使 $HA = AC$.那么(图3.26)中的某些部分以 CH 为杠杆处于平衡状态.在圆 $ABCD$ 的平面上作与 BD 平行的任一直线 MN,它与其它直线的交点如图所示,过 MN 作与 AC 垂直的平面,该平面截圆柱得一直径为 MN 的圆,截球体得一直径为 OP 的圆,截圆锥得一直径为 QR 的圆.

(a) 证明:$MS \cdot SQ = OS^2 + SQ^2$.

(b) 证明:$HA : AS = MS : SQ$.如果等式右端两项同乘以 MS,试证 $HA : AS = MS^2 : (OS^2 + SQ^2) = MN^2 : (OP^2 + QR^2)$.同时说明该比例式中最后一项等于以 MN 为直径的圆的面积与分别以 OP 和 QR 为直径的圆面积之和的比.

(c) 若以 H 为重心放置球中的圆和圆锥中的圆,那么,处于原位置上的圆柱中的圆与前两圆关于点 A 保持平衡.

(d) 由上面的结论,阿基米德得出,若把球和圆锥放在它们的重心 H 上,它们将与处于原位置上的圆柱关于点 A 保持平衡.试证 $HK : AK$ = 圆柱:(球 + 圆锥 AEF).

(e) 因为圆柱体积是圆锥 AEF 的3倍,而圆锥 AEF 的体积又是圆锥 ABD 的8倍,所以球的体积是圆锥 ABD 的4倍.

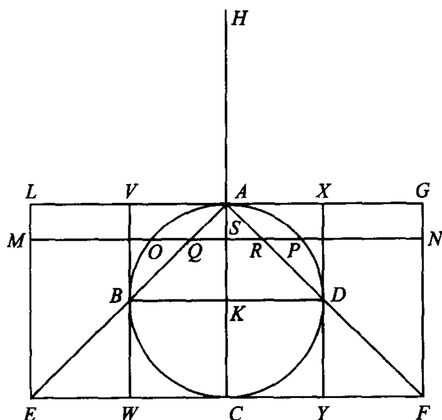


图3.26　阿基米德《方法论》命题2.

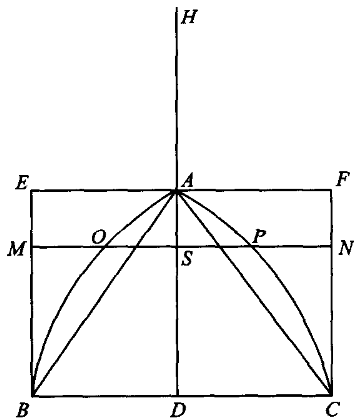


图3.27　阿基米德《方法论》命题4.

11. 命题4:旋转抛物面被垂直于轴的平面截取的部分的体积是与该部分立体同底同轴的圆锥体积的3/2倍.试用《方法论》中的方法推导这一命题.设 ABC 是抛物线 $BOAPC$ 的内接三角形(图3.27),同时三角形与抛物线都内接于矩形 $EFCB$ 中,以 AD 为轴旋转整个图形后得到一个同时内接于抛物面和圆柱的圆锥.延长 DA 到点 H,使得 $AD = AH$;作平行于 BC 的直线 MN;考虑过 MN 的平面分别截圆锥、抛物面和圆柱而得的截面,再把以 A 为中点的线段 HD 作为杠杆,用阿基米德的平衡法可证,如果以 MS 为半径的圆柱截圆放在其初始位置上,那么它与放在 H 点的以 OS 为半径的抛物面截圆保持平衡.应用圆锥的体积是它外接圆柱体积的1/3这一结论,即可证明定理.

阿基米德的其它工作

12. 用微积分证明阿基米德的下面结果，抛物线的面积是其内接三角形面积的 4/3.
13. 用分析法证明抛物弓形的顶点是抛物线上到抛物弓形底边的垂直距离最大的点.
14. 用微积分证明阿基米德的下面结果：旋转抛物面的体积是与其同底同轴的圆锥体积的 3/2.
15. 用微积分证明阿基米德的下面结果：由螺线的第一圈和初始线段所围图形的面积，等于半径为 $2\pi a$ 的圆面积的 1/3，此处螺线的极坐标方程是 $r = a\theta$.
16. 推广阿基米德计算公比为 1/4 的几何级数和的结果，计算公比为 $1/n$ 的几何级数的和.
17. 考虑《论球和圆柱》卷二中的命题 1，求与已知圆柱体积相等的球. 此处给出解决这一问题的分析过程. 设 V 是已知的圆柱，作的一个体积等于 $(3/2)V$ 的圆柱 P，再另作一个高与直径相等且体积等于 P 的圆柱 Q. 这时以 Q 的高为直径的球即为问题的解，因为该球的体积是圆柱 P 的 2/3. 于是问题转化为已知高和直径的圆柱 P，如何作直径与高相等且体积等于 P 的圆柱 Q.
18. 求抛物线和双曲线的方程，使它们的交点是三次方程 $3ax^2 - x^3 = 4a^2b$ 的根，该三次方程是阿基米德解决《论球和圆柱》卷二命题 4 所必须的. 在同一坐标系里画出两曲线的草图.
19. 证明阿基米德关于前 n 个自然数的平方和的结果可以表示为
$$\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$
20. 有一个阿基米德用旋转抛物面状的"火镜"点燃港口中敌人战船的故事，如果要设计一个可以点燃离镜子 100 米处的战船的"火镜"，那么用以旋转而得到合适的抛物面的抛物线方程是什么？这样的镜子要多大？故事中的事情有多大的真实性？

阿波罗尼乌斯对圆锥曲线的研究

21. 证明：对抛物线 $y^2 = px$，p 等于正焦弦的长度，即抛物线上过焦点与轴垂直的线段长度.
22. 把双曲线和椭圆的方程 $y^2 = x\left(p + \frac{p}{2a}x\right)$、$y^2 = x\left(p - \frac{p}{2a}x\right)$ 分别表示成现在的标准形式. 哪一点是曲线的中心？试证在椭圆情形，如果 $2b$ 是其短轴长度，那么 $b^2 = pa/2$.
23. 用微积分证明《圆锥曲线》卷 1 命题 33.
24. 用微积分证明《圆锥曲线》卷 1 命题 34.
25. 阿波罗尼乌斯叙述并证明的圆锥曲线性质远多于我们所列出的这些命题. 他的研究并不只限于诸如椭圆长轴和短轴这样的主要直径，他还研究了共轭直径. 在椭圆上任一点作一切线，过椭圆中心作与切线平行的线段，则称该线段与过切点和中心的线段共轭(图 3.28). (也可类似地定义双曲线的共轭直径，但这里只讨论椭圆的情形.)

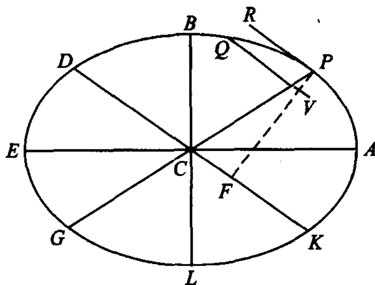


图 3.28　椭圆的共轭直径.

(a) 证明 DK 与 PG 共轭，PG 与 DK 也共轭.

(b) 设在直角坐标系中，椭圆的方程是，$b^2x^2 + a^2y^2 = a^2b^2$，$\angle PCA = \theta$，$\angle DCA = \alpha$. 已知直径 DK 与过点 $P = (x_0, y_0)$ 的切线平行，证明
$$\tan\theta = y_0/x_0, \quad \tan\alpha = -b^2x_0/a^2y_0.$$

(c) 以共轭直径 DK 和 PG 为斜坐标轴，把椭圆的方程表示成斜坐标 x', y' 的形式. 证明该变换可由下式表示：
$$x = x'\cos\theta + y'\cos\alpha,$$
$$y = x'\sin\theta + y'\sin\alpha.$$
并证明得到椭圆的新方程是 $Ax'^2 + Cy'^2 = a^2b^2$，其中

$$A = b^2\cos^2\theta + a^2\sin^2\theta,\quad C = b^2\cos^2\alpha + a^2\sin^2\alpha.$$

(d) 设 $Q=(x',y')$ 是椭圆上以 PG 和 DK 为坐标系的点，QV 是点 Q 到直径 PG 且与 DK 平行的线段，令 $QV=y'$，$PV=x'_1$，$GV=x'_2$，$PC=a'$，$DC=b'$，证明椭圆方程可以表示成 $y'^2=\frac{b'^2}{a'^2}x'_1x'_2$ 的形式.这样，可以把前面提到的阿波罗尼乌斯的命题Ⅰ—21加以推广.

(e) 证明：任何一对共轭直径构成面积相等的矩形(命题Ⅶ—31)，即若 PF 是到 DK 的垂线，则

$$PF\times CD = AC\times BC = ab.$$

26. 用《圆锥曲线》卷二命题8证明，双曲线的切线上由切点到两渐近线间的线段相等.再用非微积分的方法证明，曲线 $y=1/x$ 的切线在点$(x_0,1/x_0)$处的斜率等于 $-1/x_0^2$.

27. 已知以 C 为圆心，AA' 为直径，方程是 $y^2=x\left(p-\frac{p}{2a}x\right)$ 的椭圆，设 G 是 AA' 上满足 $AG>p/2$ 的任一点(图3.29).在 AG 上取点 N，使得 $NG:CN=p:2a$，用分析法证明，如果过 N 作与轴垂直的线 NP 交曲线于点 P，则 PG 是 G 到曲线上的最短线段；再证 PG 与曲线的过点 P 的切线垂直.

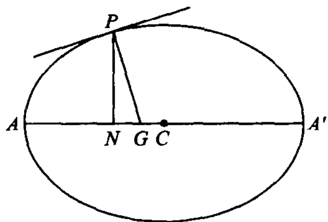


图3.29　求椭圆的垂线.

28. 改变上题的结果，求双曲线上一点的法线.

29. 怎样按所要求的个数作(用直尺和圆规)椭圆 $y^2=x\left(p-\frac{p}{2a}x\right)$ 和双曲线 $xy=k$ 上的一些点.

30. 用微积分证明，双曲线上一点到焦点的连线和该点的切线所成的角等于该切线和与轴平行的直线所成的角.

31. 用分析法证明，一条直线垂直另两条直线的三线轨迹问题的答案是圆锥曲线.设定两条平行线间的距离及已知比，画出曲线.

32. 用分析法证明，一般的三线轨迹问题的答案总是圆锥曲线.

33. 用割圆曲线三等分任意角，即已知割圆曲线 BZK 和 $\angle BAZ$，如何作 $\angle BAX$，使得 $3\angle BAX=\angle BAZ$(图3.6).

34. 把三等分角问题的下述解答补充完整(该解答由帕普斯给出，但很可能能追溯到更早的年代)[20].把已知 $\angle AOG$ 放在圆的中心，在圆上切得圆弧 AG(图3.30).要三等分这个角，只要把圆弧 AG 三等分即可，即在圆周上找一点 B，满足弧 BG 等于弧 AB 的一半.用分析法，假设 B 点已作出，那么 $\angle BGA=2\angle BAG$.作 $\angle BGA$ 的角分线 GD，另作 DE 和 BZ 垂直于 AG.应用《原本》卷六命题3和相似性证明 $BG:EZ=AG:AE=2:1$.再应用焦点－准线性质可证，B 点位于某一双曲线上，然后完成综合证明.

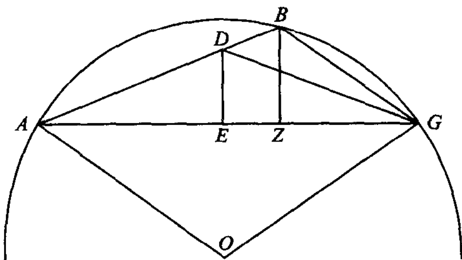


图3.30　用圆锥曲线三等分角，第二种方法.

讨论题

35. 在学习微积分前，设计一节类似阿基米德推导几何级数求和公式的课.

36. 阿基米德求抛物弓形面积和(或)螺线的一转围成图形的面积的过程是否可以被引入到微积分前(甚至微积分)课程中来计算曲线围成的图形的面积.

37. 在学习微积分前，设计一节类似阿波罗尼乌斯用圆锥的截面定义圆锥曲线，从而得出圆锥曲线方程的课，并将此方法与现代标准课本中所使用的定义作比较.

38. 在学习微积分前，设计一些论证圆锥曲线焦点和切线基本性质的课.

39. 阿波罗尼乌斯用方程对圆锥曲线的研究与现在解析几何中对此问题的论述相似程度如何？你认为他是解析几何的发明者吗？

40. 你认为阿基米德是微积分的发明者吗?

41. 为什么古希腊人要不懈地寻求化圆为方、三等分角和倍立方体问题的答案?这些答案是何用时被找到的?

文献和注解

在第2章所引用的关于希腊数学的著作中,有许多都包含有关于阿基米德和阿波罗尼乌斯工作的介绍,特别是在 Thoms Heath 的 *A History of Greek Mathematics*、Van de Waerden 的 *Science Awakening* 和 Wilbur Knorr 的 *The Ancient Tradition of Geometric Problems* 是值得进一步阅读的好资料. 在 Ivor Thoms 的 *Selections Illustrating the History of Greek Mathematics* 中可找到本章中所引用的阿基米德和阿波罗尼乌斯的著作的节选和其它相关讨论. 在 Thoms Heath 的 *The Works of Archimedes*(New York: Dover, 1953) 中包括有阿基米德现存著作的完整的并适合于现代读者的翻译. 而对阿基米德著作最详细的讨论是在 E. J. Dijksterhuis 的 *Archimedes*(Princeton: Princeton University Press, 1987) 中, E. J. Dijksterhuis 此书的新版本中包括由 Wilbur Knorr 所给出的有关阿基米德研究的新近成果目录, 其中包括 Wilbur Knorr 自己的一些成果, 如"Archimedes and the Measurement of the Circle: A New Interpretation"(*Archive for History of Exact Sciences* 15(1976), 115 - 140)、"Archimedes and Spirals: The Heuristic Background"(*Historia Mathematica* 5(1978), 43 - 75) 和"Archimedes and the Elements: Proposal for a revised Chronological ordering of the Archimedean corpus"(*Archive for History of Exact Sciences* 19(1978), 211 - 290). R. Catesby Taliaferro 在 *Great Books*(Chicago: Encyclopedia Britannica, 1952) 给出了阿波罗尼乌斯《圆锥曲线》(*Conics*) 惟一的部分(前三卷) 英文译本. Thoms Heath 的 *Apollonius of Perga*(Cambridge: W. Heffer and Sons, 1961) 包括了《圆锥曲线》现存的7卷全部内容, 由于他调整了这7卷的顺序, 同时又常常将某几个定理组合在一起, 所以这不能看作是原文献的直译, 尽管为此, 这仍然是目前惟一可以利用、完整且带有注释的阿波罗尼乌斯主要工作的英文文本.

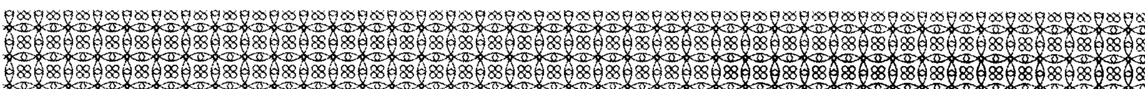
1. 引自阿波罗尼乌斯《圆锥曲线》(*Conics*) 卷1序, Heath 译, *Apollonius of Perga*, pp. lxx - lxxi.
2. Vitruvius, *On Architecture*, (Cambridge: Harvard University Press, 1934), IX, 9 - 10.
3. 在 Thoms Heath 的 *A History of Greek Mathematic*, 有对亚里士多德《机械学》(*mechanic*) 的论述, 在 Thoms 的 *Selections Illustrating the History of Greek Mathematics* 中可找到这些章节的译文.
4. Plutarch, *The Lives the Noble Crecians and Romans*(Dryden translation), 载于 *Great Books*, 14, p. 254. 引自论 Marcellus 的章节(下同).
5. Plutarch, *The Lives*, p. 252.
6. E. J. Dijksterhuis, *Archimedes*, p. 299. 这里引用的是 Pappus: Collectio Ⅷ, 5, 11 的翻译.
7. Plutarch, *The Lives*, p. 252.
8. 同上, p. 253.
9. 同上, p. 253.
10. 王冠问题的讨论引自 Heath 的 *The Works of Archimedes*, pp. 259 - 260. Heath 深入地介绍了阿基米德的各种数学技巧.
11. Heath 给出了阿基米德《方法论》的翻译, 并把它作为上述著作的一个附录. 这里引自其中的第13页. 在 Asger Aaboe 的 *Episodes from the Early History of Mathematics*(Washington, MAA, 1964) 中出可找到有关这方面的有价值的讨论, 而 S. H. Gould 的"The Method of Archimedes,"(*American Mathematical Monthly* 62(1955), 473 - 476) 一文中则给出了《方法论》的一个简述.
12. 同上, p. 17.
13. 同上.
14. 对希腊数学中不可分量的使用的讨论可参见 Wilbur Knorr 的"The Method of Indivisibles in Ancient Geometry," 载于 Ronald Calinger(ed.), *Vita Mathematica*(Washington MAA, 1996), pp. 67 - 86.
15. Livy, *History of Rome*(Cambridge: Harvard University Press, 1934), XXIV, sec. 34.
16. 这里的图片和讨论引自 Knorr 的 *The Ancient Tradition of Geometric Problems*, Knorr 在此书中就阿波罗尼乌斯对希腊

几何问题求解所作的贡献作了广泛的讨论.

17. 引自阿波罗尼乌斯《圆锥曲线》的前三卷,转引自 R. Catesby Taliaferro 的译本: *Great Books*.
18. 此定理出现在 Thoms Heath 的 *Apollonius of Perga* 第 143 和 152 页.
19. 对 Diocles 工作的讨论引自 Gerald Toomer 的 *Diocles on Burning Mirrors*(New York: Springer, 1976),书中包括有完整的翻译并且讨论了 Diocles 工作的重要性.
20. Knorr, *The Ancient Tradition*, p. 128.

阿基米德、阿波罗尼乌斯及其先驱者的工作概览

公元前 5 世纪	希波克拉底(Hippocrates)	倍立方体问题
公元前 4 世纪中期	梅内赫莫斯(Menaechmus)	首次对圆锥进行切割
公元前 4 世纪后期	阿里斯托斯(Aristaeus)	圆锥曲线的早期著作
公元前 287— 前 212	阿基米德(Archimedes)	数学模型、面积和体积
公元前 3 世纪后期	尼科米迪斯(Nicomedes)	曲线和化圆为方
公元前 250— 前 175	阿波罗尼乌斯(Apollonius)	圆锥曲线
公元前 2 世纪早期	狄俄克利斯(Diocles)	火镜



第4章　古希腊时代的数学方法

> 柏拉图……给数学家们提出了以下的问题：应该假定行星是作什么样的匀速而完全有规律的圆周运动，才能使这些星球所呈现出来的表观运动得以保持？
>
> 辛普利修斯(Simplicius)："评亚里士多德的《论天体》"

现征聘：急需计算工作者从事繁重但是例行的计算，以编制天文学主要工作所需的表格．应聘者应能准确地接受详尽的指示．报酬：包吃、包住外加未来1200年内将要利用这些表格的成千上万人的一片感激之情．联系人：克劳蒂乌斯・托勒密(Claudius Ptolemy)，地址：天文台(这是一则刊登在大约公元前150年的亚历山大时期报纸上的分类广告)．

虽然并没有真正出现过这种广告，克劳蒂乌斯・托勒密倒真是写了一本重要的著作来回应柏拉图的挑战，这是一本一直被人们学习、研究、评注和受到广泛批评的著作，也是一本长达一千四百年之久从未被取代过的著作，在这本著作中托勒密不仅运用了来自平面几何和球面几何的许多概念，而且发明了一些方法来开展繁杂的数值计算，这种计算为使他的著作成为一本非常有用的书是不可缺少的．托勒密的这本教材，以及其他一些来自巴比伦和埃及的古代天文学著作，在占星学中被广泛地采用．但尽管如此，来自所有这些文明的迹象都表明，研究天文学的最主要的目的是为了解决与历法有关的问题——例如季节确定，日蚀的预测，以及太阴月起始日的确定等等．

在运用数学来研究天文学的过程中，希腊人创立了平面三角学和球面三角学，而且还创建了宇宙的一个数学模型．从柏拉图时代到托勒密时代，在这长达五个世纪的时间里，他们对这个模型作过多次修正．在数学天文学的发展过程中作出过杰出贡献的人有很多，本章将对他们的思想一一加以讨论，这些人中包括有：公元前4世纪的欧多克斯(Eudoxus)，第3世纪末的阿波罗尼奥斯(Apollonius)，第2世纪的喜帕恰斯(Hipparchus)，公元100年左右的梅内劳斯(Menelaus)，以及最后还有托勒密．在本章的最后部分将概述一下在古希腊发展起来的"实用数学"的工作，这就是指用来解决地球上而不是有关天体问题的数学．特别地，我们将讨论公元第1世纪海伦(Heron)以及第2世纪拉比・尼希米亚(Rabbi Nehemiah)所做的工作．

4.1　托勒密之前的天文学

通过几个世纪对天体的观察,巴比伦人已经能够对许多重复发生的天体现象作出相对来说比较准确的预测.这些预测包括从简单的(例如日出、日落时间等等),到相当复杂的(例如月蚀的时间).他们用于这些研究的技巧最多也就是一些算术和简单的代数而已.毕达哥拉斯学派也用过数来解释一些天文现象.但无论是巴比伦人还是毕达哥拉斯学派,他们都从未建立过一个能够将各种天体现象联系起来的模型.直到柏拉图学派时代,即在第4世纪的希腊,才首次创建了一个这样的模型.

他们所建立的基本模型是由两个同心球组成的,一个代表地球;另一个代表其他星体.虽然来自我们感觉的直接迹象显示地球是平的,但稍微深入一点的观察却让希腊人相信地球是球形的.在他们的观察中包括了这样一些事实:当一艘船从开出时,船体总是比船的桅杆的顶端先从视线中消失,以及当月蚀发生时,地球在月球上的阴影的边缘是圆形的.在他们的美学观念中,球被认为是最完美的实体形象,这一美学观更加坚定了他们的信念.同时很自然地,他们认为其他天体的形状应该和地球的形状一样.

直觉的证据和逻辑的论证,进一步让希腊人相信,地球是静止不动的,并且处于天体的中心.这后一个结论是根据各主要天体现象的对称性推断出来的,而前一个结论则是从人们没有感觉到地球的运动这一点推断出来的.希腊人还提出了如下的论点:"假如地球每天沿自身的轴旋转一周,那么它的速度一定非常快,使得任何不停在地球表面上的物体看上去都在作与地球转动方向相反的同样的运动,无论是天上的云彩还是其它飞行物或是被抛出的物体,我们永远不会看到它们向东运动,这是因为地球自转引起的向东的速度总是会赶超过这些物体运动的速度,而使这些物体看上去总是在向西,向后运动"[2].在认为地球是固定不动的这一前提下,每天所观察到的天空上的各种运动就必然是由于天球的转动引起的,而所有的恒星都是牢牢地连接在天球上,一组一组地分布形成的图案就叫做星座图.这些恒星间的位置是从来不变的,并形成了"漫游星"或行星的固定背景(见补遗4.1).

补遗 4.1	哥白尼的先行者
	对于文中讨论过的认为地球不动的"地心说",有些古代天文学家就提出过与之对立的理论.小亚细亚蓬托斯(Pontus)地方的赫拉克里德斯(Heraclides)(约388 B.C—310 B.C)曾断言,日常看到的天体运动是来自地球的自转,而根据阿基米德的报导,萨莫斯(Samos)的阿里斯塔丘斯(Aristarchus)(约310 B.C—230 B.C.)还提出了这样的假说:"恒星和太阳维持不动,而地球则围绕着太阳作圆周运动,太阳处于轨道的中心"[3].反对阿里斯塔丘斯的一个主要理由是,它意味着恒星的表观位置将随着地球在其轨道上的位置的不同而不断改变.针对这个反对意见,阿里斯塔丘斯进一步假设,地球与这些恒星之间的距离是如此巨大,以致于上面所说的变化是觉察不到的.那时其他天文学家不相信可能有这么大的距离.另外有些思想家指控他没有信仰,因为他为了解释天体现象而"不惜让整个宇宙动起来"[4].科学与宗教的冲突显然早在古代就开始了.

七大游星 —— 太阳、月亮、水星、火星、金星、木星加土星 —— 则是比较松驰地连接在天球上.这些游星附着在天球是很显然的;从整体来看,它们每天都参与天球由东往西的旋转运动.但是它

们同时又有自身的运动,通常是往相反的方向,即从西往东,而且速度也要慢得多.正是这种运动是希腊天文学家(事实上包括所有早期的天文学家)所力图理解的.然而希腊人的努力受到了他们下述压倒一切的哲学立场的局限:假如地球以外的宇宙真如亚里士多德所说的那样是完美和永恒不变的,那么天体的运动就只能是这些完美物体的"自然"运动.由于这些物体是球形的,其自然运动就是圆周运动.于是,天文学家和数学家就试图通过用匀速率的圆周运动的各种几何构造的组合来建立一个模型,从而解释天体中的各种现象("拯救表观现象")—— 即试图解决在本章开头所引述的柏拉图的问题.尽管确定这些运动是否可能和物理上如何来实现并不是这些天文学家、数学家们的事 —— 我们知道在古希腊从来不研究天体物理,但事实上他们成功地找出了能对付柏拉图的挑战的几种体系.

因为希腊人的基本天体模型是由球体组成,所以研究天体运动的首要的要素就是研究球的性质.我们记得,欧几里得的《原本》几乎没有提到这些性质.不过写于第4世纪一些别的教本,包括比坦诺(Pitane)地方的奥托利可斯(Autolycui)(约公元前300年)的著作以及欧几里得本人的其它著作倒是涉及到了球面的一些基本知识,其中大部份结果都是天文学中直接用得上的.这些书中包括了像**大圆**(通过球心的平面与球的截面)和**极点**(与大圆所在平面垂直的直径的两个端点)这样一些概念的定义.这些教本还包含了三个重要的定理:(1)球面上任何不在同一直径上的两点惟一决定一个大圆.(2)任何一个大圆,如果它通过第二个大圆的极点,那么它必与第二个大圆相互垂直,并且这第二个大圆也一定通过第一个大圆的极点.(3)任二大圆必相互等分.



图4.1　一件以色列纪念品上的黄道十二宫镶嵌图.

在天球上有几个大圆对天文学来说是非常重要的.例如,太阳从西往东通过各星球的轨道是一个大圆,这个大圆叫做**黄道**,它通过黄道带的十二个星座(图4.1)(这些星座首先是在巴比伦的天文学中提及,并早在公元前300年的希腊文献中出现).通过地球南北极的直径,向两端延长到天穹,就是天体每日绕其旋转的转轴.与这条轴上的极点相对应的大圆就叫做**天球赤道**,这条赤道与黄道相交于一条直径的两个端点,这两个端点就分别叫**春分**和**秋分**,因为在这两天,太阳就处于这两个交点上(图4.2).黄道上离赤道最远的两个点,一个在赤道的北面,一个在赤道的南面,这两个

点则分别叫做夏至和冬至.

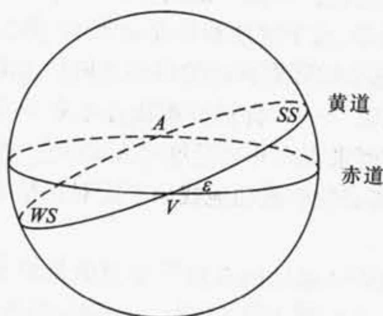


图 4.2 黄道与赤道(V.春分,A 秋分,SS 夏至,WS 冬至,N 北极点).

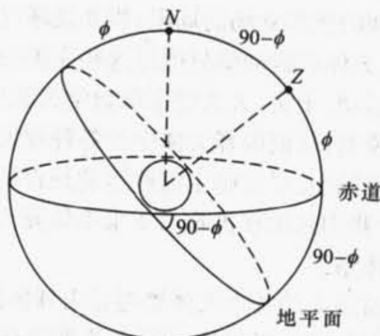


图 4.3 地平面与赤道.

由于希腊人知道,相对于布满星体的球面来说,地球是如此之小,以致于它实际上可以看成是一个点,因此他们假定地平面通过天球的中心,从而地平面本身也就是一个大圆.地平面与赤道分别在东、西两点相交.最后本地子午线是指通过地平面对应的南、北极点和头顶正上方的本地天顶点的大圆.由于这个子午圆同时与地平面和天球赤道二者相垂直,因此,它必定经过天球赤道对应的南、北极点.黄道与赤道的交角 ε 可以这样来确定:取太阳在夏至和冬至这两天正午时分的地平纬度之差的一半.在欧几里得时代,其测得值为 24°,而在托勒密时代其测得值为 23°51′20″(事实上,这个值一直在慢慢地减小,如今它的值大约为 $23\frac{1}{2}$°).地平面与赤道平面的交角是 $90° - \varphi$,这里 φ 是观察者所在处的地理纬度(图 4.3).

4.1.1 欧多克斯和多球模型

以他在比率和穷竭法方面的工作而著称的欧多克斯是一位将天文学转化为数学的主要功臣.他很可能就是这个二球模型的发明人,并对模型作了许多必要的修改以便于说明太阳、月亮和恒星的各种运动.但不管怎样改,他还是谨遵柏拉图的要求,只用圆周运动.在他的模型中,所有的天体都被置于一组相互关联的球的内球面上,这组球有两个或两个以上,都以地球为中心,用它们绕不同的轴的转动来产生所观察到的运动(图 4.4).例如,太阳需要用两个球才能说明它的基本运动.外球为太阳以外的各星星所在的球,它每天沿着自己的轴由东往西转一周.太阳所在的内球则连在外球上,它的转轴与外球的转轴成一夹角 ε.现在,如果让内球沿着自身的轴由西往东,并且以每年一周的速度旋转,那么内外球两种运动的叠加就产生了太阳的表观运动(图 4.5).对于月亮则需要用三个球来说明.外球还是同样的每天由东往西转一周,最里面的球则由西往东 $27\frac{1}{3}$ 天转一周,这里的 $27\frac{1}{3}$ 天是月亮绕黄道一周所需的时间.但由于月亮的轨道与黄道之间有多达 5° 的偏离,欧多克斯用了一个中间球,这个球与外球成 ε 角的倾斜,而与内球则成 5° 角的倾斜,它的由东向西的缓慢运动至少可以定性地产生月球在南北极之间的偏差.为了能描述更为复杂的行星运



图 4.4 利比里亚邮票,其左下角为欧多克斯的多球模型(所有球以地球为中心).

动,包括像从整体看来由西向东运动、但偶然地又有逆行(即由东向西)这样的现象,欧多克斯就要用四个球[5].

在所有这些可能的模型中,欧多克斯认为这些球都是用来辅助计算的工具,而不是实实在在的实体.虽然可以找出具体的参数数值,使模型系统能表示多种天体运动,但这些模型系统并不能解释所有观察到的自然现象.例如,行星四球理论并不能预测它们在运行过程中亮度的明显变化.尽管如此,亚里士多德还是把经过改进了的欧多克斯的多球模型作为物理实体,并把它纳入到他那详尽的宇宙学里面了.就这样,这一理论成了天体基本概念中的一部份,在西方文明史上,这些基本概念一直持续到16世纪.

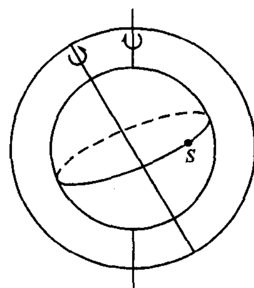


图 4.5 用来说明太阳运动的欧多克斯两球模型.

4.1.2 阿波罗尼奥斯:偏心圆和本轮

大约在欧克多斯后150年,阿波罗尼奥斯试图对柏拉图的问题找出新的解答.在那时之前的200年人们就已经知道一年四季并不都是等长的.例如,从春分到夏至的时间要比从夏至到秋分的时间长两天.因此,认为太阳是围绕以地球为中心的圆周作匀速运动这样一个简单模型,就算认为太阳是附着在欧多克斯球上的这样的模型,都无法解释这样的现象.因为非匀速的运动不符合柏拉图的规则,因此阿波罗尼奥斯或他的前人提出如下解决方案,即把太阳轨道的中心置于离地球有一定距离的地方(叫做**偏心点**).这样一来,如果太阳沿着这个新的轨道(叫做**均轮**)作匀速运动,那么在地球上的观察者就会看到春季象限(右上)超过四分之一圆周的量比夏季象限(左上)所超过的量要大(图4.6(a))距离 ED,或者更准确地讲,是 ED 与 DS 的比值,叫做这个均轮的**偏心率**.如将直线 ED 延长到均轮,那么离地球最近的交点就叫做均轮的**近地点**(perigee),而离地球最远的交点就叫做**远地点**(apogee).假定人们可以确定这个模型中的正确参数值(即 ED 的长度和方向),而使得由此所得的季节长度与实际观测所得一致,则运用这个模型要解决的问题就是,太阳每一天出现在什么地方.为了解决这个问题,我们需要求出角 DES.这就需要去解三角形 DES,即需要三角学的知识.事实上,正是由于在这个几何模型中引起一些数值参数,才导致了三角学的创建.

阿波罗尼奥斯还注意到上了上述偏心圆模型可以用另一个几何模型来取代,这个模型就是本轮.就是说,不再认为太阳是循着偏心圆运动,取而代之,我们设想太阳在一个叫**本轮**的小圆上运动,而这个圆的圆心则沿原来的以地球为中心的圆周运动(图4.6b).如果本轮在其圆心绕地球一周的同时也沿顺时针方向自转一周 —— 也就是说,如果这两种运动总是保持 $DECS$ 为平行四边形

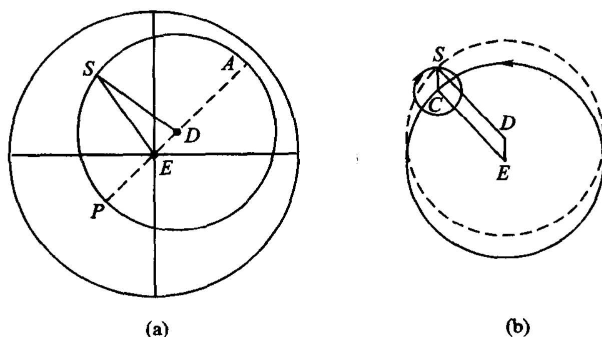


图 4.6 (a) 阿波罗尼奥斯用来说明太阳运动的偏心圆模型;(b) 阿波罗尼奥斯用于太阳的本轮模型.

—— 那么太阳的实际轨道就会和上面用均轮解释的一样.

于是人们就发现,如果将偏心圆与本轮相结合,则会得到行星更复杂的运动.阿波罗尼奥斯已实际上发动了对这一模型的研究.图(4.7a)表示一行星在一中心为 C 的本轮上逆时针匀速运动,点 C 则在以 D 为中心(D 距地球远为 DE)的均轮上沿同一方向运动.如果沿这些圆周运动的速度恰当,那么从地球上将可以观察到:从总体上来说该行星将沿着黄道向东运动,但在某些特定时间将会向相反的方向运动(当行星处于本轮的里侧时)(图 4.7b).为了利用这一模型,同样也需要求出各有关的参数值,例如 PC 和 ED 的长度和它们的相对方向.然而,一旦对某特定的行星确立了这些参数值,我们就可以通过解一定的三角形来求出该行星在任何时刻的位置.

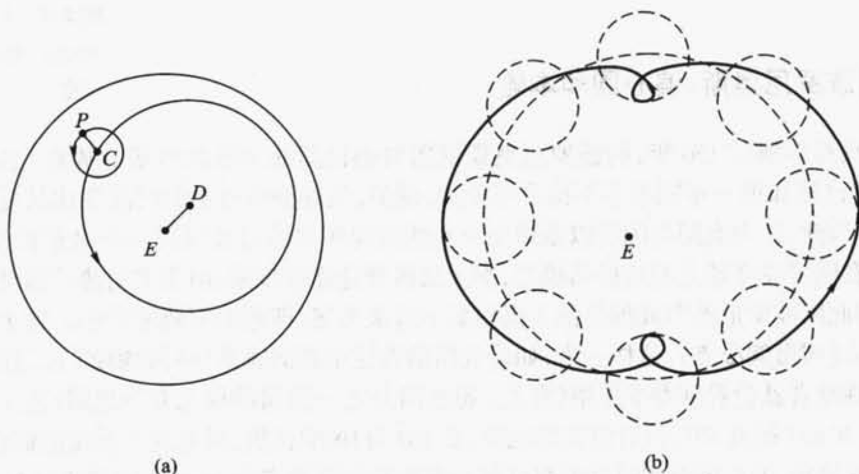


图 4.7 行星运动的阿波罗尼奥斯模型.

4.1.3 希帕科斯和三角学的兴起

阿波罗尼奥斯本人并不具备为解决上一节中提出的问题而必需的三角学知识.是系统地做了大量有关行星位置观察的比西尼亚(Bithynia)地方的希帕科斯(Hipparchus)(190 B.C—120 B.C),给星球引进了坐标系统,并开始编制一些有关三角比率的表格,这些表使得天文学家能够很容易地求解直角三角形和成功地处理阿波罗尼奥斯的问题(图 4.8).为了定量地处理恒星和行星的位置,我们不仅需要知道弧长和角度的量度单位,同时还需要有一种方法来确定某一物体在天球上所处的位置 —— 即需要一个坐标系.欧几里得的角度单位只不过是直角,其它的角度则是用这一角度的分数或倍数来描述.不过,早在公元前 300 年以前巴比伦人就创始将圆周等分成 360 个部份,叫做度.在那以后的两个世纪中,这一度量法,并随同度的六十进制分割成的分和秒一起为希腊世界所采用.希帕科斯是最早使用这种度量单位的人之一,虽然他也用过一个圆周的 1/24 弧长和一个圆周的 1/48 弧长,在他的著作中把这叫做"步(Steps)"和"半步(half-steps)".为什么巴比伦人要把圆周分成360等分,我们尚不得而知.也许是因为 360 这个数容易被许多小的数字整除,或者是因为360是最接近一年中天数的整数值.这后一个理由给了我们一种很方



图 4.8 一张希腊邮票上的希帕科斯.

便的求太阳每天沿黄道走过多少度的近似方法,这个近似值刚好是 1.

最早把坐标引进天穹的也是巴比伦人,他们所采用的系统叫黄道坐标系,后来为托勒密所利用.在这个坐标系中,恒星的位置由沿着黄道和垂直黄道所测得的距离来确定.沿黄道的坐标值(从北极看从春分点逆时针方向走过的角度)叫**经度** λ;垂直的坐标值(以度为单位所表示的在黄道北面或南面离开黄道的距离)就叫做**纬度** β(图 4.9a).这个坐标系在处理太阳、月亮和行星运动时特别有用.由于太阳沿黄道运动,因此它的纬度总是0°,而它的经度则从春分时的0°开始,每天增加约1°,到夏至时为 90°,秋分时为 180°,以及到冬至时为 270°.不过,不管是在巴比伦的文献还是在希腊资料中,经度常常是用十二宫图的记号来计算的;也就是说,将黄道等分成 12 个区段,每段为 30°,分别用十二宫图中的星座来命名.例如,纬度从 0° 到 30° 这一段叫白羊座(Aries),从 30° 到 60° 这段叫金牛座(Taurus).因此,如果说太阳的经度是金牛座 5°,其实就是说它的黄道经度是 35°.

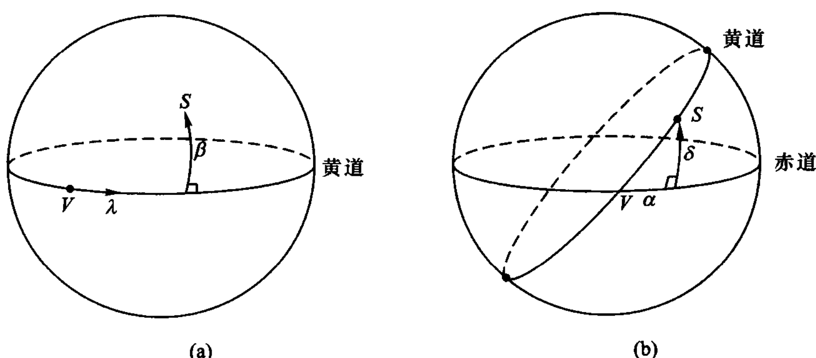


图 4.9 (a) 在天球上的黄道坐标系;(b) 在天球上的赤道坐标系.

希帕科斯还用了基于天球赤道的坐标系来取代这一黄道坐标系.沿赤道方向的坐标(也是从春分开始沿逆时针方向走过的度数)叫**赤经**(right ascension)α.其垂直方向的坐标(从赤道起向北或向南离开的距离)叫**赤纬**(declination)δ(图 4.9b).在希帕科斯制订的恒星目录中,对有些恒星位置的描述就用了这个坐标系.

为了能将一个坐标系中一个点的坐标与它在另一坐标系中的坐标联系起来 —— 这对于解决天文问题是必要的 —— 这就需要球面三角学.但要发展球面三角学,必须首先掌握平面三角学.希帕科斯显然是第一个试图将长度列成详细的表格,以便能来求解平面三角形的问题.尽管还没有直接的文献载有希帕科斯的表格或他的方法,但是来自各种资料的信息已足以使我们能对他的工作勾勒出一幅合理的图画.

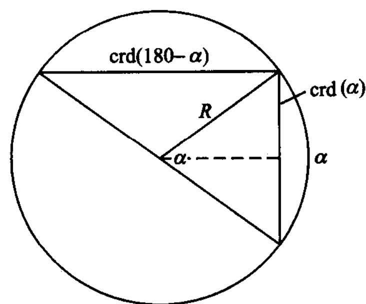


图 4.10 chord(α) 和 chord(180 − α).

在希帕科斯(以及后来的托勒密)的三角学中,一个基本元素就是在一半径固定的圆中给定弧(或中心角)所对的弦.他们两人都对各种不同的弧长 α,列出了对应的弦长 chord(α) 的值表.注意,这里 chord(α),以后编写为 crd(α),是一个长度(图 4.10).如果圆的半径记为 R,则 α 角所对的弦与其正弦之间有下述关系

$$\frac{1}{2}\mathrm{crd}(\alpha)/R = \sin\frac{\alpha}{2} \quad 或 \quad \mathrm{crd}(\alpha) = 2R\sin\frac{\alpha}{2}.$$

因为角度或弧度是用度或分来度量的,所以希帕科斯决定对圆半径也用同样的度量.已知圆周长为 $2\pi R$,取 π 的六十进制近似值 3;8,30(此值接近于阿基米德值 $3\frac{10}{21}$ 和 $3\frac{1}{7}$ 的平均值),他算得近似到最接近整数值的半径 R 的度数为 $\frac{60\cdot 360}{2\pi} = \frac{6,0,0}{6;17} = 57,18 = 3438'$.在这个半径值的圆中,一个角的度量(定义为其对应的圆弧长除以圆的半径)就等于它的弧度的度量.

为了算出对弦的表,希帕科斯从 60° 角开始算起.在这种情况下,对弦长就等于半径值,即 $\mathrm{crd}(60°) = 3438' = 57,18$.对 90° 角,对弦长等于 $R\sqrt{2} = 4862' = 81,2$(注意我们在这里将十进制记法和六十进制记法混合使用,这在希腊和现代的角度度量中是常见的).为计算其他角度的对弦,希帕科斯利用了两个几何结果.首先由图 4.10 可见,$\mathrm{crd}(180-\alpha) = \sqrt{(2R)^2 - \mathrm{crd}^2(\alpha)}$,因为 $\mathrm{crd}(180-\alpha) = 2R\cos\frac{\alpha}{2}$,这一结果就相当于 $\sin^2\alpha + \cos^2\alpha = 1$.其次,希帕科斯利用了一种半角公式算出了 $\mathrm{crd}\left(\frac{\alpha}{2}\right)$(人们猜测他是用了后来由托勒密给出的同样的办法).假设角 $\alpha = \angle BOC$ 为 OD 平分(图 4.11).为了用 $\mathrm{crd}(\alpha) = BC$ 来表达 $\mathrm{crd}\left(\frac{\alpha}{2}\right) = DC$,在 AC 上选一点 E,使得 $AE = AB$.那么 $\triangle ABD$ 与 $\triangle AED$ 全等,从而 $BD = DE$.由于 $BD = DC$,因而 $DC = DE$.如果从 D 作 EC 的垂线 DF,那么 $CF = \frac{1}{2}CE = \frac{1}{2}(AC - AE) = \frac{1}{2}(AC - AB) = \frac{1}{2}(2R - \mathrm{crd}(180-\alpha))$.但由于三角形 ACD 与 DCF 相似,因而 $AC : CD = CD : CF$.因此

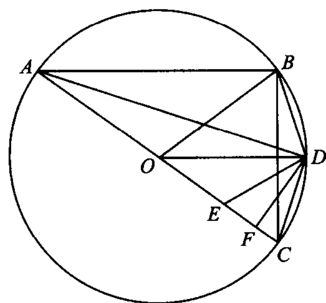


图 4.11　希帕科斯 - 托勒密半角公式.

$$\mathrm{crd}^2\left(\frac{\alpha}{2}\right) = CD^2 = AC\cdot CF = R(2R - \mathrm{crd}(180-\alpha)).$$

将上式用现代的记号来表示就得

$$\left(2R\sin\frac{\alpha}{4}\right)^2 = R\left(2R - 2R\cos\frac{\alpha}{2}\right),$$

或者,如用 2α 代替 α,就得

$$\sin^2\frac{\alpha}{2} = \frac{1-\cos\alpha}{2},$$

这正是标准的半角公式.

于是,希帕科斯就轻而易举地算出了从 $7\frac{1}{2}°$ 开始到 180° 为止、以 $7\frac{1}{2}°$ 为"半步长"的任一角的对弦值.例如,对 crd(60°) 三次应用上面的半角公式,就算得 $\mathrm{crd}\left(7\frac{1}{2}°\right)$.再用补角公式就得到 $\mathrm{crd}\left(172\frac{1}{2}°\right)$.这样算出来的数量有限的表格已经能够使希帕科斯在求解三角形的问题上取得一些进展,并使他能利用这些结果来完成他的天体模型.不过由于希帕科斯的原著已经遗失,因此我们必需将目光转向古代最有影响的天文著作,那就是克劳蒂乌斯 · 托勒密(Claudius Ptolemy)的天文学《大成》.

4.2　托勒密与《大成》

对克劳蒂乌斯·托勒密(公元年约100—178年)的个人生涯,人们一无所知,只知道他在亚历山大港做过大量的天文观察,并写了几本重要的书(图 4.12). 例如《地理》一书记录了世界上已知的地方和它们对应的经度与纬度,其中还讨论了为制作地图所需要的投影知识. 他还写过天文、音乐和光学方面的著作,也曾试图对欧几里得的平行公设给出证明. 不过,在今天他最享盛名的著作还是《数学大全》(Mathematiki Syntaxis),共 13 卷,书中对希腊人的宇宙模型给出了完整的数学描述,并包括有太阳、月亮和行星的各种运动参数.《数学大全》一书是古希腊天文学的光辉顶点,就像欧几里得的《原本》一样,它取代了这一课题的所有早期的著作. 从此书问世开始到16世纪,它一直是最有影响力的天文著作,并经历了无数次的翻印、分析和研究. 天文学家能够创造出数学模型 —— 即对自然现象的一种定量描述 —— 来可靠地预测,对于这一点,没有任何一本书能比《数学大全》更具有震撼力. 不管是在伊斯兰国家,还是在西方国家,直至包括哥白尼的工作,实际上几乎所有后来的天文工作都是建立在托勒密的这本杰作基础之上的. 在此书出版后的许多个世纪之后,这本书被誉为**伟大的文集**(megisti Syntaxis),以示它与那些平庸的天文著作的区别. 后来伊斯兰的科学家开始改称此书为 al-magisti(大成),从此本书就以《大成》(Almagest) 而闻名于世(图 4.13).



图 4.12　托勒密(头上戴着皇冠,手里拿着地球仪)(选自拉斐尔的油画《雅典学派》的细部. 头上的皇冠表明拉斐尔将托勒密与埃及的国王混为一谈).

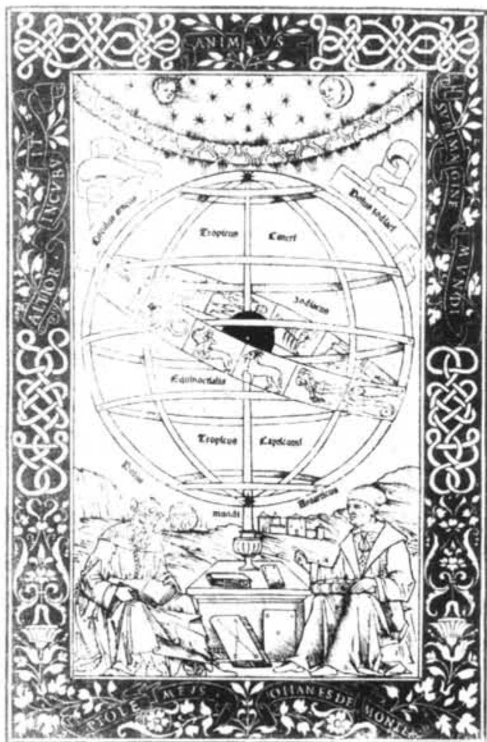


图 4.13　一本早期出版(1496) 的《大成》摘要中的一幅版画. (来源:Smithson 研究所图书馆,Photo No.76 – 14409).

4.2.1 弦　表

在《大成》一书的一开始，托勒密首先对古希腊人的宇宙概念作了简单的介绍，紧接着就给出了计算行星位置所必需的，有关平面三角学与球面三角学的数学材料.托勒密的第一个任务就是要构造出一张比希帕科斯的更完整的对弦表.为了构造从$\frac{1}{2}°$到180°，以$\frac{1}{2}°$为间隔的弦表，并找出一种能在已算好的两个值之间的插值方法，他需要有比希帕科斯更多的几何知识.同时考虑到原来的$R = 57,18$在计算中很不方便，他采用了$R = 60$，这是六十进制中的一个单位值，而托勒密的全部计算都是在六十进制中进行的.

托勒密首先计算了36°的对弦，即圆内接十边形的边长.如图4.14，ADC是以D为圆心的圆的直径.BD垂直于ADC，E平分DC，取点F使得$EF = EB$.根据《原本》Ⅱ－6，我们有$CF \cdot FD + ED^2 = EF^2 = BE^2$，* 因此$CF \cdot FD = BE^2 - ED^2 = BD^2 = CD^2$，从而$D$以黄金分割比分割$CF$.又根据《原本》ⅩⅢ－9，如果将同一个圆的圆内接正六边形和正十边形的边排成一直线，则其交接点分割这条直线段成黄金分割比.因为半径CD与圆内接正六边形的边长相等，因此托勒密证明了DF是正十边形的边长，也就是说$DF = \text{crd}(36°)$.为了算出其数值，他注意到

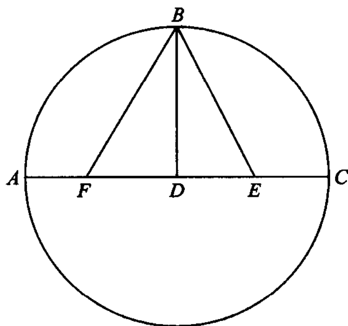


图 4.14　托勒密对 crd(36°) 的计算.

$$DF = EF - ED = EB - ED = \sqrt{BD^2 + ED^2} - ED = \sqrt{3600 + 900} - 30 = 37;4,55.$$

接下来托勒密又注意到，由于正五边形的边长$C = \text{crd}(72°)$的平方等于正十边形的边长的平方与正六边形边长的平方之和(《原本》ⅩⅢ－10)，因此$\text{crd}(72°) = \sqrt{R^2 + \text{crd}^2(36°)} = 70;32,3$.这里我们自然用了$\text{crd}(60°) = 60$，此外，$\text{crd}(90°) = \sqrt{2R^2} = \sqrt{7200} = 84;51,10$以及$\text{crd}(120°) = \sqrt{3R^2} = 103;55,23$.最后，由于$\text{crd}^2(180 - \alpha) = (2R)^2 - \text{crd}^2\alpha$，因而对于任何一个已知其对弦值的角度，托勒密同时也能算出它的补角的对弦.例如，$\text{crd}(144°) = 114;7,37$.因此，仅仅从欧几里得几何的一些命题和平方根的计算出发，他为构造弦表开了一个好头.

就像四个世纪之前的阿基米德一样，托勒密从未讲过他是怎样去计算平方根的，只是直接把结果摆出来.在第4世纪末塞翁(Theon)在有关托勒密的工作的一个评注中，提出了一种可能是托勒密一直在用的方法："如果想求任何一个数的平方根，那么先找出一个最接近的整数平方根值，将它乘以2，然后将(原来的数减去平方根近似整数的平方所得)余数换成"分"值，再除以上面的乘积，(所得商的整数就得到了"分"位的值)再减去商的平方；依次类推，再将余数换成"秒"值，又除以已得平方根近似到度和分的近似值的2倍(所得商的整数部份就是平方根的"秒"位的值).这样我就得到了正方形的一条边(即平方根)的所要求的近似值."[6]**

塞翁提出的这种用来计算近似到六十进制中两位小数点的求根方法完全类似于我们在第一章中所讨论过的中国人的求根算法.例如要计算$\sqrt{7200}$，首先注意到$84^2 = 7056$且$85^2 = 7225$，因而答案一定是$84;x,y$的形式.由于$7200 - 84^2 = 144$，将$144 \cdot 60$(将余数化成分值)除以$2 \cdot 84$，得商的近似整数值为51.于是我们知道答案为$84;51,y$的形式.最后，$7200 - (84;51)^2 = 0;28,29$，将它约化

* 原文误作 $CF \cdot FD + ED^2 = EF^2 \cdot BE^2$—— 译者注.

** 本段原文过于简略，为便于读者理解，添加了括号内的文字.请读者对照下一段的实例 —— 译者注.

为秒值就是 1719. 把它除以 $2\cdot 84;51 = 169;42$,得其整数近似值为 10. 因此所要求的平方根的近似值为 84;51,10. 这一计算相当复杂,而托勒密对大量的这种计算又都只摆出其结果,因此人们认为托勒密一定是依靠了大量"计算员"来从事这种冗长乏味却又是不可缺少的工作. 特别地,利用上面已经算出了的基本弦值和希帕科斯的半角公式,以及根据下面一个新定理而推出来的和角与差角公式,托勒密完成了他的弦表,而在此过程中,"计算员"的帮助是不可或缺的.

托勒密定理　给定一个圆内接四边形,它的对角线的乘积等于它的两对边乘积之和.

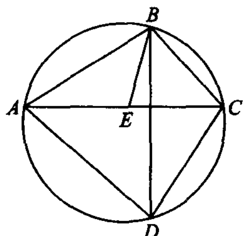


图 4.15　托勒密定理.

如图4.15,在四边形 $ABCD$ 中,需要证明 $AC\cdot BD = AB\cdot CD + AD\cdot BC$. 在 AC 上取点 E 使得 $\angle ABE = \angle DBC$,那么 $\angle ABD = \angle EBC$. 且由于 $\angle BDA$ 与 $\angle BCA$ 所对的圆弧相同,所以还有 $\angle BDA = \angle BCA$. 因而 ΔABD 与 ΔEBC 相似. 于是 $BD : AD = BC : EC$,或者说 $AD\cdot BC = BD\cdot EC$. 类似地我们有 $\angle BAC = \angle BDC$,ΔABE 与 ΔDBC 相似,因而 $AB : AE = BD : CD$,或 $AB\cdot CD = BD\cdot AE$. 将这两个等式相加就得 $AB\cdot CD + AD\cdot BC = BD\cdot AE + BD\cdot EC = BD(AE + EC) = BD\cdot AC$,定理由此得证.

为了推导两条弧长 α,β 之差的对弦公式,托勒密在上述定理中令 $AC = \text{crd}(\alpha)$(图 4.16),将定理的结论应用于四边形 $ABCD$,有 $AB\cdot CD + AD\cdot BC = AC\cdot BD$(图 4.16),由于 $BC = \text{crd}(\alpha-\beta)$,

$$120\ \text{crd}(\alpha-\beta) = \text{crd}\ \alpha\cdot\text{crd}(180-\beta) - \text{crd}\ \beta\cdot\text{crd}(180-\alpha).$$

这很容易翻译成现代的差角正弦公式:

$$\sin(\alpha-\beta) = \sin\alpha\cdot\cos\beta - \cos\alpha\cdot\sin\beta.$$

类似地可以证明有

$$120\ \text{crd}(180-(\alpha+\beta)) = \text{crd}(180-\alpha)\text{crd}(180-\beta) - \text{crd}\ \beta\cdot\text{crd}\ \alpha,$$

与这个公式等价的是下述的和角余弦公式

$$\cos(\alpha+\beta) = \cos\alpha\cdot\cos\beta - \sin\alpha\cdot\sin\beta.$$

运用差角公式和半角公式,托勒密接着就算出了 $\text{crd}(12°) = \text{crd}(72° - 60°)$,$\text{crd}(6°) = \text{crd}\left(\frac{1}{2}\cdot 12°\right)$,$\text{crd}(3°)$,$\text{crd}\left(1\frac{1}{2}°\right)$ 及 $\text{crd}\left(\frac{3}{4}°\right)$. 他算出的最后两个结果分别为 $\text{crd}\left(1\frac{1}{2}°\right) = 1;34,15$ 和 $\text{crd}\left(\frac{3}{4}°\right) = 0;47,8$. 利用和角公式,他现在已经能建立起步长为 $1\frac{1}{2}°$,甚至是 $\frac{3}{4}°$ 的弦表. 不过由于他所需要的是步长为 $\frac{1}{2}°$ 的列表,而且,由于"给定了像 $1\frac{1}{2}°$ 的对弦值,那么其对应的三分之一弧长的对弦是无法用几何方法求出的(如果可解的话,我们就立刻可以得到 $\frac{1}{2}°$ 的对弦了)",托勒密只得用特定的办法来求 $\text{crd}(1°)$ 和 $\text{crd}\left(\frac{1}{2}°\right)$,此法"虽然一般来说并不能准确地确定出对弦的值,但在弧长很小的情况下,却能得到误差小到可以忽略不计的近似值"[7]. 换言之,虽然托勒密并未给出证明,但他坚信,欧几里得工具(即几何方法)不足以用来确定 $\text{crd}\left(\frac{1}{2}°\right)$,或者,更一般地来说,是不能将一个角三等分的. 因此,必需采用别的方法.

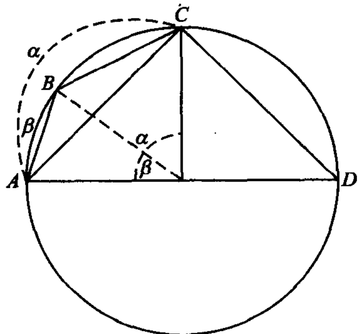


图 4.16　差角的对弦公式.

这种另类方法其实就是一种近似的方法,它是根据如下的一个引理:如果 $\alpha < \beta$,则 $\text{crd}\beta : \text{crd}\ \alpha < \beta : \alpha$,或者,用现代的术

语来表达，这就是说$\frac{\sin x}{x}$随着x趋于 0 而增加. 先将这个引理应用到$\alpha = \frac{3}{4}°, \beta = 1°$，托勒密得到$\mathrm{crd}(1°) < \frac{4}{3}\mathrm{crd}\left(\frac{3}{4}°\right) = \frac{3}{4}(0;47,8) = 1;2,50,40$. 然后再将引理应用到$\alpha = 1°, \beta = 1\frac{1}{2}°$，他得到了$\mathrm{crd}(1°) > \frac{2}{3}\mathrm{crd}\left(1\frac{1}{2}°\right) = \frac{2}{3}(1;34,15) = 1;2,50$. 因为一切的计算都是近似到六十进制的小数点后的第二位，看来这时我们有$\mathrm{crd}(1°) = 1;2,50$，从而$\mathrm{crd}\left(\frac{1}{2}°\right) = 0;31,25$. 再利用和角公式，托勒密就可以构造出从$\mathrm{crd}\left(\frac{1}{2}°\right)$到$\mathrm{crd}(180°)$，步长为$\frac{1}{2}°$的弦表了. 为了有助于计算任何"分"值弧长的对弦的插值，他在他的表中附加了第三列，这一列给出了从$\mathrm{crd}(\alpha)$到$\mathrm{crd}\left(\alpha + \frac{1}{2}°\right)$的增加值的三十分之一. 下面是他的表格中的一小部分，其精度大致相当于现代十进制小数点后第五位.

弧长	对弦	六十分之一	弧长	对弦	六十分之一
$\frac{1}{2}$	0;31,25	0;1,2,50	6	6;16,49	0;1,2,44
1	1;2,50	0;1,2,50	47	47;51,0	0;0,57,34
$1\frac{1}{2}$	1;34,15	0;1,2,50	49	49;45,48	0;0,57,7
2	2;5,40	0;1,2,50	72	70;32,3	0;0,50,45
$2\frac{1}{2}$	2;37,4	0;1,2,48	80	77;8,5	0;0,48,3
3	3;8,28	0;1,2,48	108	97;4,56	0;0,36,50
4	4;11,16	0;1,2,47	120	103;55,23	0;0,31,18
$4\frac{1}{2}$	4;42,40	0;1,2,47	133	110;2,50	0;0,24,56

4.2.2 解平面三角形

有了弦表以后，托勒密已经可以开始解平面三角形的问题了. 虽然他从未叙述过有关解决这类问题的系统方法，但似乎他确实采用了确定的规则. 当把托勒密的方法与现代的方法作比较时，必须牢记一个重要的差别，那就是在托勒密的弦表中，对弦是指圆半径为 60 时，弦的长度，而不是一个比率值. 因此，在给定的实际问题中，他总是需要针对实际的半径值去调整表中的数值. 在此我们通过三个具体的例子来看他的解题方法.

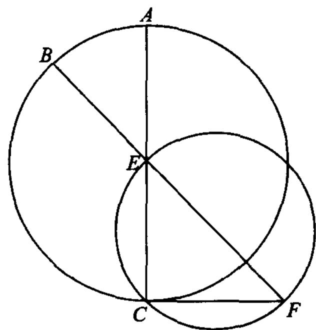


图 4.17 计算柱的阴影长度.

第一个例子，计算罗得岛(Rhodes，北纬 36°) 上一根长度为 60 的柱子 CE 在春分日中午时分的日影长 CF，托勒密首先注意到当时太阳是在天顶点下 36°(也就是说$\angle AEB = 36°$)(图 4.17). 托勒密把 CF 当作三角形 ECF 的外接圆的一根弦. 由于同一圆弧所对的圆心角

是圆周角的二倍,所以 $CF = \text{crd}(72°) = 70;32,3$,同时 $CE = \text{crd}(180° - 72°) = \text{crd}(108°) = 97;4,56$. 由于托勒密需要的是当 $CE = 60$ 时日影的长度,所以他将所算出的值按比率$\frac{60}{97;4,50}$缩小,因此所要求的日影长为$\frac{60}{97;4,58} \cdot (70;32,3) = 43;36$,这个问题其实是解一直角三角形,已知该三角形的一边 b 和角 α,求另一边 a 的长,于是上面的计算可重新写为:

$$a = b \cdot \frac{\text{crd}(2\alpha)}{\text{crd}(180 - 2\alpha)} = b \cdot \frac{2R\sin\alpha}{2R\cos\alpha} = b \cdot \tan\alpha,$$

这和现代的方法是一致的.正是由于托勒密在当时缺少像正切函数这样的概念,以及他总是要用一个具体圆的弦来计算,这就迫使他不但要计算两倍给定角的对弦值,同时还要计算两倍给定角的补角的对弦值以及它们的商.

第二个例子表明了,托勒密是怎样来计算太阳偏心圆模型中的各个参数的.[8] 这个计算相当于要解直角三角形 LDE,这里 D 为太阳轨道的圆心,E 为地球(图 4.18,比较图 4.6a).用通过点 E 的两条相互垂直的直线将黄道分割成四个象限,并且类似地把偏心圆也分成四个象限.为了计算 LD 和 LE 的长度,必须先利用已知的各季节的长度来计算弧长 $\theta = \frac{1}{2}\widehat{VV'}$ 和 $\tau = \frac{1}{2}\widehat{WW'}$.已知春季太阳的轨道长为 94.5 天,而夏季的太阳轨道长为 92.5 天,再设太阳平均每天角速度为 ν,那么由图可见:对春季来说有 $90 + \theta + \tau = 94.5\nu$,而对夏季来说则有 $90 + \theta - \tau = 92.5\nu$.由于 ν 等于用一年的天数(当时的观察值为 365;14,48)去除 360°,简单的计算就得 $\theta = 2°10'$ 和 $\tau = 0°59'$.

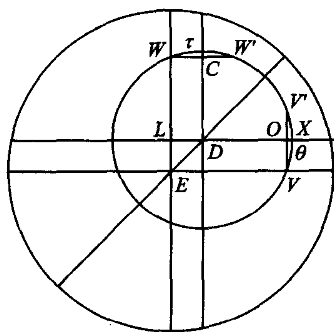


图 4.18 太阳偏心圆模型的参数计算.

如果设均轮圆的半径 DX 长为 60,则三角形 DLE 的边长就可以算出了.由于 DX 平分弧 VV',显然 $LE = OV = \frac{1}{2}VV' = \frac{1}{2}\text{crd}\,2\theta = \frac{1}{2}\text{crd}(4°20°) = 2;16$.类似地,$DL = \frac{1}{2}\text{crd}2\tau = \frac{1}{2}\text{crd}(1°58°) = 1;2$.根据毕达哥拉斯定理则有 $DE^2 = LE^2 + DL^2 = 6;12,20$,从而 $DE = 2;29,30$,或近似于 $2;30 = 2\frac{1}{2}$.用现代术语来讲,托勒密仅仅只是计算了 $LE = OV = R\sin\theta$,以及 $DL = CW = R\sin\tau$.由于经常需要计算两倍角的对弦的一半,这就导致后来的天文学家编制了这个量的表格,这就是现代的正弦函数表.

为了完成本例中的三角问题的解,托勒密用了 ΔLDE 的外接圆来计算 $\angle LED$.由于当 $DE = 2;29,30$ 时 $LD = 1;2$,因此,如果 $DE = 120$,LD 则等于 49;46.现在反过来使用弦表,托勒密就得到了它所对应的弧长为 49°.由此可知,$\angle LED$ 等于它的一半,即 24°30'.因而 $\angle LDE = 65°30'$,这个三角形由此得解,同样地,用现代的术语来讲,托勒密是首先计算了 $120a/c = 2R\sin\alpha$,或 $\sin\alpha = a/c$,然后再用反正弦关系来确定 α 的值.

最后是托勒密提供的解一个斜三角形的例子.问题是这样的,在偏心圆模型中,任意取 DS 的长为 60,当给定 $DE = 2;30$ 时,求出太阳的方向角 $\angle DES$(图 4.19).对任何给定的一天,利用太阳在其轨道上的速度就能算出角 PDS,从而也能得知角 EDS.托勒密在他的计

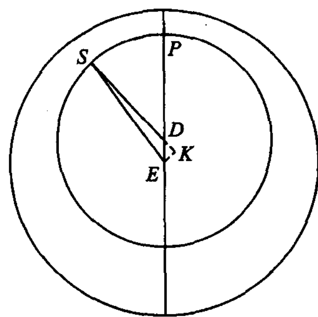


图 4.19 求太阳的位置.

算中取 $\angle PDS = 30°$ 和 $\angle EDS = 150°$. 他首先从点 E 向 SD 的延长线作垂线 EK. 和以前一样, 考虑绕三角形 DKE 的圆, 得弧长 $DK = 120°$. 他从弦表中注意到, 如果半径为 60(或 $DE = 120$), 则 DK 就等于 crd(120°) = 103;55. 然而, 由于 $DE = 2;30$, 则按比例应有 $DK = 2;10$. 因此, $SK = SD + DK = 62;10$. 由于 $\angle KDE = 30°$, 且 $EK = \frac{1}{2}DE = 1;15$. 应用毕达哥拉斯定理到 ΔSKE 上, 就得到 $SE = 62;11$. 接着考虑 ΔSKE 的外接圆, 由于当 $SE = 62;11$ 时 $KE = 1;15$, 所以, 如果 $SE = 120$, 则 KE 会等于 2;25. 反查弦表, 就可以求得弦长 2;25 所对应的弧为 2°18′, 由此得知 $\angle KSE = 1°9'$, 因而 $\angle DES$ 等于 $180° - 150° - 1°9' = 28°51'$.

这个托勒密的方法可翻译如下: ΔABC 中, a, b 已给, 且 $\gamma > 90°$, AD 为 BC 延长线的垂线(图 4.20). 如 $AD = h$, $CD = p$, 则

$$p = \frac{\mathrm{crd}(2\gamma - 180) \cdot b}{2R}, \quad h = \frac{\mathrm{crd}(360 - 2\gamma) \cdot b}{2R},$$

由此可得

$$\begin{aligned} c^2 &= h^2 + (a + p)^2 \\ &= a^2 + \left(\frac{\mathrm{crd}^2(360 - 2\gamma)}{4R^2} + \frac{\mathrm{crd}^2(2\gamma - 180)}{4R^2}\right) b^2 + \frac{2ab\,\mathrm{crd}(2\gamma - 180)}{2R} \\ &= a^2 + b^2 + 2ab\,\frac{\mathrm{crd}(2\gamma - 180)}{2R}, \end{aligned}$$

或

$$c^2 = a^2 + b^2 - 2ab\cos\gamma.$$

这正是当三角形的两边长及其夹角为已知时, 求第三边长的余弦定律, 为了求这个三角形的其他角, 托勒密注意到

$$\mathrm{crd}(2\beta) = \frac{h \cdot 2R}{c},$$

从而由弦表可求出 β. 这可翻译成

$$\sin\beta = \frac{h}{c} = \frac{b \cdot \sin\gamma}{c}.$$

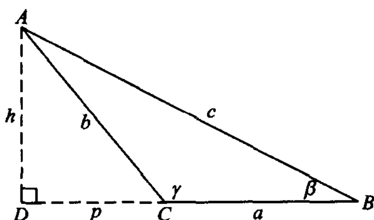


图 4.20　托勒密的余弦定理.

可见托勒密也使用了正弦定律的等价形式.

应当指出的是, 在上面这一例中托勒密明显地给出了由给定 a, b 和 γ 的值来计算 c 和 β 的一个算法. 实际上, 类似这样的算法在《大成》中是随处可见的. 这些平面三角学的算法完全可以不折不扣地翻译成现代的三角学公式.

4.2.3　解球面三角形

托勒密甚至还广泛地讨论了解球面三角形的算法. 虽然早在公元前 300 年就已有对球面几何的研究工作, 但最早的球面三角学的著作却是梅内劳斯(Menelaus)的《球面》一书(约公元前 100 年). 这本书中的一个主要的结果就是我们今天所谓的梅内劳斯定理, 它给出了如图 4.21 所示的球面上一组大圆的圆弧之间的关系. 两圆弧 AB 和 AC 为另两圆弧 BE, CD 所切割, 而 BE, CD 相交于点 F. 各圆弧的标记如图所示, 并设 $AB = m$, $AC = n$, $CD = s$ 和 $BE = r$, 如果用正弦而不是用对弦来表示, 梅内劳斯定理就是说

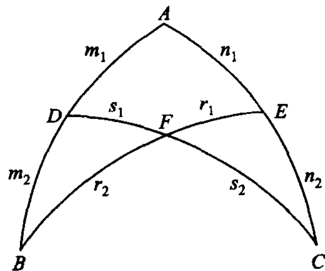


图 4.21　球面上的梅内劳斯位形.

$$\frac{\sin(n_2)}{\sin(n_1)}=\frac{\sin(s_2)}{\sin(s_1)}\cdot\frac{\sin(m_2)}{\sin(m)},\qquad(4.1)$$

以及

$$\frac{\sin(n)}{\sin(n_1)}=\frac{\sin(s)}{\sin(s_1)}\cdot\frac{\sin(r_2)}{\sin(r)}.\qquad(4.2)$$

梅内劳斯证明了这些结论(《大成》一书中也有相同的证明),证明过程是这样的:首先在类似的平面位形中证明这些结论,然后再将球面上的图形投影到一个平面上.而托勒密就用梅内劳斯定理来解球面直角三角形,这就是由大圆的圆弧组成的三角形,其中有两条弧相交成直角.设有这样的一个三角形,其直角顶点在 C,角 C,B,A 所对的边记为 c,b,a(图 4.22),托勒密构造了包含这样一个三角形的梅内劳斯位形.例如,如果 ABC 为一正角三角形,分别作大圆 PM 和 QN,它们各以 A,B 为其极点,然后再将三角形的每条边延长,使之与这两个大圆相交.这样就得到了两组梅内劳斯位形,一个是以 M 为顶点,另一个以 N 为顶点.由于大圆上一段弧的长度与该弧对该大圆极点所张角的度数相等,同时由于 P 和 Q 分别为 QM 和 PN 的极点,因而定理中的方程就可以被大大简化而用来得到所给三角形的边与角之间的关系.

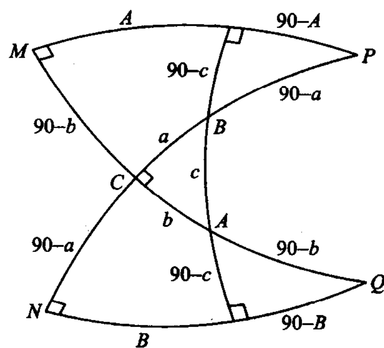


图 4.22 托勒密构造的双梅内劳斯位形.

首先,如果应用以 M 为顶点的位形,方程(4.1)就变成

$$\frac{\sin(90-A)}{\sin A}=\frac{\sin(90-a)}{\sin a}\cdot\frac{\sin b}{\sin 90}\quad\text{或}\quad\tan A=\frac{\tan a}{\sin b},\qquad(4.3)$$

方程(4.2)则成为

$$\frac{\sin 90}{\sin A}=\frac{\sin 90}{\sin a}\cdot\frac{\sin c}{\sin 90}\quad\text{或}\quad\sin A=\frac{\sin a}{\sin c}.\qquad(4.4)$$

其次,如果应用以 N 为顶点的位形,方程(4.1)就成了

$$\frac{\sin a}{\sin(90-a)}=\frac{\sin c}{\sin(90-c)}\cdot\frac{\sin(90-B)}{\sin 90}\quad\text{或}\quad\cos B=\frac{\tan a}{\tan c},\qquad(4.5)$$

而方程(4.2)则成了

$$\frac{\sin 90}{\sin(90-a)}=\frac{\sin 90}{\sin(90-c)}\cdot\frac{\sin(90-b)}{\sin 90}\quad\text{或}\quad\cos c=\cos a\cdot\cos b.\qquad(4.6)$$

托勒密对这些结果的第一个应用是:已知经度 λ,求太阳的赤纬 δ 和赤经 α(图 4.23).图中 V 为春分点,VA 是赤道,VB 是黄道.根据托勒密已算出的结果,黄道与赤道的交角 ε 为 $23°51'20''$.假设太阳位于经度为 λ 的点 H.为了求出 $HC=\delta$ 和 $VC=\alpha$,必须求解球面三角形 VHC.由方程(4.4),$\sin\varepsilon=\sin\delta/\sin\lambda$,或 $\sin\delta=\sin\varepsilon\cdot\sin\lambda$.托勒密分别对 $\lambda=30°$ 或 $\lambda=60°$ 两种情况进行了计算.得出在前者情况下 $\delta=11°4'$,而在后者情况下 $\delta=20°30'9''$.他向他的计算人员这样演示了他的算法之后,就让他们去进行计算,对从 $1°$ 到 $90°$ 每一个给定的整数 λ 编制 δ 的数值表.类似地,由方程(4.5),$\cos\varepsilon=\tan\alpha/\tan\lambda$,或 $\tan\alpha=\cos\varepsilon\tan\lambda$.再次,托勒密另外对 $\lambda=30°$ 算得

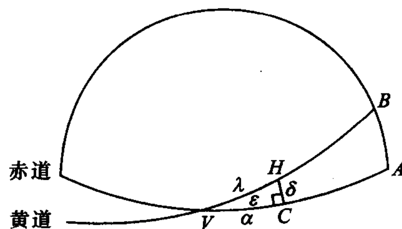


图 4.23 给定太阳的经度,计算其赤经和赤纬的方法.

α 为 $27°50'$，而对 $\lambda = 60°$ 算得 α 为 $57°44'$. 然后他列出了 α 相对其它 λ 的数值. 另外注意到，根据对称性，还有 $\alpha(\lambda + 180) = \alpha(\lambda) + 180°$，以及 $\delta(\lambda + 180) = -\delta(\lambda)$.

托勒密所解决的很多其它问题都是与确定一段黄道圆弧的"日出时间"(rising time) 紧密相关的. 也就是说，在一给定的地理纬度，托勒密需要确定天球赤道上的一段弧长，它与给定的黄道上，一段圆弧同一时间与地平线相交. 不失一般性，可以假设要求圆弧的一个端点为春分点，因此只要求出赤道上 VE 的长度，这 VE 与黄道上的给定圆弧 VH 同时与地平线相交(图 4.24). 这段弧长就叫做"日出时间"，这是因为时间是由赤道绕其轴的匀速运动来度量的. 转一整周的时间是 24 小时，所以赤道上的 15° 就相当于 1 个小时，而 1° 则相当于 4 分钟. 在任何情况下，要解决托勒密问题，只需要解三角形 HCE，求出 $EC = G(\lambda, \phi)$，然后用已经确定了的 $VC = \alpha(\lambda)$ 来减去这个值就可以了. 具体举例来说，假设纬度为 $\phi = 36°$，且 $\lambda = 30°$. 根据上面的计算，$\delta = 11°40'$. 由方程(4.3) 就得出 $\sin G = \tan\delta/\tan(90 - \phi) = \tan\delta\tan\phi$，从而有 $G = 8°38'$. 由于 $\alpha = 27°50'$，日出时间 $VE = 27°50' - 8°38' = 19°12'$. 托勒密(或他的工作集体) 对从 10° 到 360°，步长为 10° 的所有 λ 值，以及对十一个不同的纬度值 ϕ，分别计算了日出时间 $\rho(\lambda, \phi)$，并将结果编成一张很长的表.

现在就可以用这张表来计算在任一给定纬度上在任何一天的白昼长度 $L(\lambda, \phi)$. 由于太阳在经度为 λ 处的日落时刻就是经度为 $\lambda + 180$ 地方的日出时刻，因此我们只需将 $\lambda + 180$ 及 λ 所对应的日出时间相减就可以了. 又由于有 $\sigma(\lambda + 180, \phi) = -\sigma(\lambda, \phi)$，我们还可以作进一步的简化，有

$$\begin{aligned} L(\lambda, \phi) &= \rho(\lambda + 180, \phi) - \rho(\lambda, \phi) \\ &= \alpha(\lambda + 180) - \sigma(\lambda + 180, \phi) - \alpha(\lambda) + \sigma(\lambda, \phi) \\ &= 180° + 2\sigma(\lambda, \phi). \end{aligned}$$

例如，当 $\phi = 36°$ 和 $\lambda = 30°$ 时，$L(30, 36) = 180° + 2\sigma(30, 36) = 180° + 17°16' = 197°16'$，这大致相当于 13 小时零 9 分.

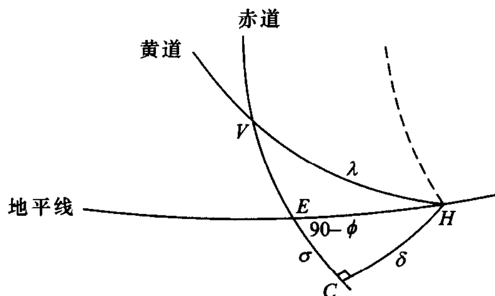


图 4.24　日出时间的计算.

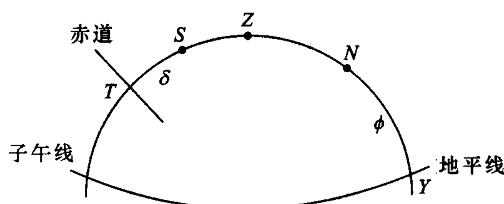


图 4.25　计算太阳离开天顶的距离.

利用图 4.24，我们还可以计算出日出的位置，即弧长 $EH = \beta$. 当纬度为 36°，$\lambda = 30°$ 时来计算此值，我们利用方程(4.4) 得

$$\sin\beta = \frac{\sin\delta}{\sin(90 - \phi)} = \frac{\sin 11°40'}{\sin 54°} = 0.25,$$

故 $\beta = 14°28'30''$. 因此，在 $\lambda = 30°$ 的这一天，太阳会在本地时间上午 5 : 25 从正东偏北 $14°28'30''$ 的地平线上升起.

作为球面三角学的最后一个应用，我们来计算中午时分太阳离开天顶的距离. 在任何一天中，太阳离开赤道的距离都是 δ，因此当太阳在中午穿过子午线时，它与子午线的交点一定是在北极 N 和子午线与赤道的交点 T 之间(假设 $\delta > 0$)，并且离该交点的距离为 δ(图 4.25). 由于弧 $NT = 90°$，

$NY = \phi$，因而弧$\widehat{SZ} = 90° - (90° - \phi) - \delta = \phi - \delta$，因此，如果$\phi - \delta > 0$，或$\phi > \delta$，那么太阳在正午时将在南边，而阴影则朝北边．由于$\delta$的最大值为23°51′20″，因此对于纬度大于这一数值的地方来说，中午时的情况就总是如上面所说的那样．另一方面，当$\phi = \delta$时，太阳正午就在头顶的正上方．对于给定的纬度，出现这种情况的日子都可以容易地算出来，总之，给定从天顶到太阳的角距离．托勒密就能计算出日影的长度，这在前面已经讲到过了．他把结果编成了一张长长的表格．对39个不同的纬度处最长白昼的长度以及一根长为60的立柱在春分、秋分以及在夏至和冬至的那几天的日影长度他都一一算了出来．

我们这里所举的例子都只是针对太阳，并且都是取自《大成》的前三卷．在该书的其余部分，托勒密还讨论了月亮和行星．对每一个天体，他首先对要解释的现象作一简短的定性描述，然后对设定的由本轮和偏心圆组合而成的几何模型作一说明，最后是根据他本人的亲自观测或者是根据他以前已经记载的观测结果，详尽地推导出该模型的各个参数．一般来说，作为结束他总是要说明，用他的模型配合所算出的参数可以预测行星的新的位置，而这个位置又可以通过观测来检验．因此，托勒密是第一位有史料为证，利用数学模型来真正"做"科学的数学科学家．他从一个模型开始，然后用观测的结果对它进行改进，直到能在他的观测精度范围内预测所观察的现象为止．

补遗 4.2

托勒密与函数的概念

作为一部数学著作，托勒密的《大成》向我们提出了这样一个问题：我们是否能够从中看到现代函数概念的萌芽？支持有这个概念的是这样一个事实，就是书中的确有许多显示几组量之间的函数关系的表格．比这早得多的巴比伦人就已经编制了平方根表、倒数表以及一些天文用表，这些表可以用来预测各种天体现象的出现时间．但是总体来说，他们只对离散值感兴趣．托勒密迈开了巨大的一步，他不但编制了各种表，而且还展示了如何进行插值以便能对"独立变量"的任何值算出对应的函数值，从而为连续现象的计算处理奠定了基础．例如，弦长是用α的函数$\mathrm{crd}(\alpha)$来表示，太阳的赤纬用经度的函数$\delta(\lambda)$来表示，而日出时间则用沿黄道的弧长λ和地理纬度ϕ这两个变量的函数$\rho(\lambda, \phi)$来表示．托勒密还经常反过来用他的表，例如对弦值来求弧长，因此，他实质上是用了我们今天所谓的反函数．

此外，由于托勒密工作的总的目的是预测行星的位置，他经常给出具体的算法，讲述如何针对任一给定时间来做到这一点．例如，要计算太阳在任一时刻的位置，托勒密讲述了需要怎样做的各个步骤．首先计算从历元(所有计算的起始时间——公元前747年2月26日)到给定时刻所经历的时间t；接着从"平均运动"表中查出平均运动$\mu(t)$；再把$\mu(t)$加上265°15′又减去360°的倍数，使其值$\bar{\lambda}$小于360°；然后再从太阳的反常表(这张表中有一项在托勒密求解斜三角形的例子中曾计算过)中查出$\bar{\lambda}$对应的值$\theta(\bar{\lambda})$；然后再将$\theta(\bar{\lambda})$与$\bar{\lambda}$和65°30′相加，就得到最后的结果．用现代的记号来表示就是：$p(t) = \theta(\bar{\lambda}(t)) + \bar{\lambda}(t) + 65°30' \pmod{360°}$，其中$\bar{\lambda}(t) \equiv \mu(t) + 265°15' \pmod{360°}$，而其中的$\theta$和$\mu$本身则由函数关系导出的表中查得．尽管托勒密并未使用过这些现代的记号，但他是意识到了现代的函数关系这个概念的．在他的许多算法中，他甚至用了一些合适的对称性来简化他的计算．

然而，托勒密并没有讨论函数的一般概念．事实上，他显然认为用函数来处理的方法是理所当然的．由此可以推断，这些方法很可能为他的读者所熟悉，并且早已有人用过，至少在他之前的天文学家就用过了．虽然如此，没有任何证据能表明古希腊数学家在函数这个题目上写过什么东西，这也许是由于没有什么好的理论工具来处理函数及其性质的缘故．也没有与之相关的公设．但是有必要注意到在"官方希腊数学的几何外衣"[10]背后，还存在着实用数学的领域，即是为了解决包括天上和地下两方面问题所必需的数学．

托勒密为他在"拯救表观现象"上所取得的成就而感到自豪.也就是说,他证明了,对所有七大游荡的天体来说,"它们表观上的反常运动都可以用匀速圆周运动来表示,因为只有这种运动才符合造物的神意…….因此,理所当然应该承认,为此目的而取得的成功是一件伟大的将举,它真正标志着理论哲学的数学部分的巅峰.但是,有很多原因又使我们不得不承认这是一件非常艰难的事,而且我们有充分的理由说明为什么在我们之前始终无人能够成功地做到这一点"[11].然而,托勒密克服了种种困难,并且为后代留下了一本杰出的数学著作,一本的的确确预言了天体现象的著作.这本著作在随后的 1400 年中从未被其它的著作超越过(补遗 4.2).

4.3　实用数学

希帕科斯和托勒密的三角学使得希腊人能够"测量"天上的以及地球上与天体现象有关的三角形.但希腊人想必也需要解地球上的一般三角形,以便能够对距离、高度等进行非直接测量.也许人们会很自然地认为,至少在希帕科斯之后,他们应该会利用三角学的方法,即与弦表有关的方法了.然而现有的历史资料却显示他们并没有这样做.

自然,在希帕科斯之前,人们只有直接从相似概念推导出来的间接测量方法.这正是可以从欧几里得的《光学》一书中所能找到的方法.这部论著基本上是一部有关视觉几何原理的著作,其基本假设就是,光线总是沿着直线传播的.但其中的确包括了欧几里得的几个有关间接测量的结果.例如,命题 18 是要"求一高度的大小,使得太阳不被挡住"[12].换言之,假定太阳在点 Γ,欧几里得想要算出一铁塔 AB 的高度,使得它的阴影长度为 $B\Delta$(图 4.26).先将高度 EZ 为已知的物体放在这样的路径上,使得它的阴影也以 Δ 为终点,从而它的阴影为 $E\Delta$.根据三角形 $AB\Delta$ 和 $ZE\Delta$ 的相似性,欧几里得得出了 AB 的高度.

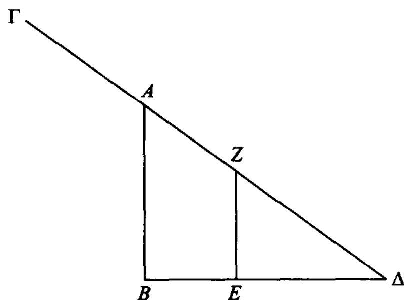


图 4.26　利用太阳来计算高度,摘自欧几里得的《光学》.

4.3.1　海伦的著作

在欧几里得之后大约 350 年,出现了有关间接测量方面的更具体的著作,那就是亚历山大地方的海伦(Heron)(公元第 1 世纪)所著的《屈光学》(屈光器是一种助视仪器).尽管从他的另一本书中可以看出他熟悉弦表,海伦在本书中还是用了相似三角形.例如,海伦为了显示如何来确定从观察者(在点 A)至一个抵达不到的点 B 之间的距离,他首先取点 Γ,使得 ΓAB 在同一直线上,然后过 Γ 作 ΓAB 的垂线 ΓE,最后作 E 到 B 的视线,从而在 BE 上可找到一点 Δ,使得 $A\Delta$ 也垂直于 $BA\Gamma$(图 4.27).由于三角形 $AB\Delta$ 与 ΓBE 相似,我们有 $\Gamma E : A\Delta = \Gamma B : BA$.头一个比值是已知的,因为它的每个长度都可以测量得到.从而等式右边的比例也可知.但 $\Gamma B : BA = (\Gamma A + AB) : BA = \Gamma A : BA + 1$,由于 ΓA 为已知,BA 因此也就可以被算出.

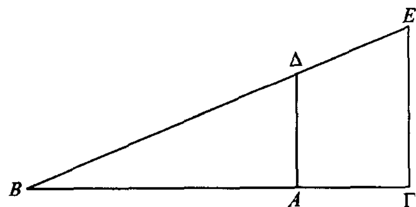


图 4.27　计算距离 BA,摘自海伦的屈光学.

海伦用了类似的方法来计算诸如两个不可抵达的点之

间的距离,诸如铁塔的高度(不用它的阴影的长度),山谷的深浅这样一些问题.他还讲述了,当从一座山的两端来开挖一条隧道时,应如何来确定两端各自的方向,以便挖出的隧道成一条直线.

海伦的许多著作还包含了应用数学中其它一些重要的概念.他的《反射光学》中就包含了一个关于当光线射到镜面上时入射角等于反射角的有趣的证明.虽然这个结果早就知道了,但他是以"自然从不作无用功"[13]这样的假设来作为他的证明的基础的,也就是说,他假定光线从物体 C 出发经过镜面反射到达眼睛点 D 的路径一定是最短距离.设 A 是镜面 GE 上的一点使得 $\angle CAE = \angle DAG$(图 4.28).分别作 DA 和 CE 的延长线,相交于点 F.则容易推出 ΔAEF 与 ΔAEC 为全等三角形,从而光程 $DA + AC$ 就等于直线段 DAF.现在假设 B 是镜面上任一其它点,连接 BF, BD 和 BC.由于 $BF = BC$,我们有 $DB + BC = DB + BF > DAF$.因此,任何其他设定的光程都会比具有入射角等于反射角的光线路程长.

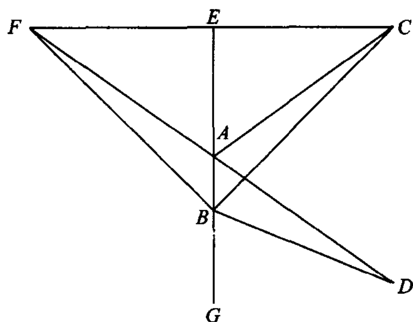


图 4.28 入射角等于反射角.

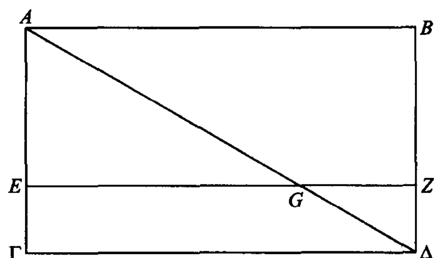


图 4.29 速度的平面四边形法则,摘自海伦的《力学》.

在海伦的《力学》一书中出现了相当于我们现在所熟悉的速度平行四边形法则(虽然这个概念在大家认为是亚里士多德所写的力学等著作中早就有了).设有一点沿直线 AB 作匀速运动,由 A 运动到 B,而与此同时直线 AB 又以匀速平行于自身运动,最后停在直线 $\Gamma\Delta$ 的位置(图4.29).设 EZ 是 AB 在运动过程中的某一中间位置,而 G 为它上面的一个动点的位置,则 $AE : A\Gamma = EG : EZ$(根据运动的定义),从而 $AE : EG = A\Gamma : EZ = A\Gamma : \Gamma\Delta$,因而 G 在对角线 $A\Delta$ 上.换句话说,对角线 $A\Delta$ 就是移动点的实际路径.如果用现代术语来说,那就是"速度向量" $\overrightarrow{A\Delta}$ 是速度向量 $\overrightarrow{AB}$ 和 $\overrightarrow{A\Gamma}$ 的向量和.

自然,希腊人自己并不曾想到过"速度向量".他们并没有把速度当作一个可以测量的独立量来考虑,也没有"每小时英哩数"这样的概念.回想起《原本》V,定义3,比率只能在同一类量的大小之间相比,从而不能考虑距离与时间的比率.人们只能在距离之间或时间之间来进行比较.因此,奥托利可斯(Autolycus)给出的一个早期的速度的定义,它是这样的:"当一个点在相等的时间里经过相等的距离时,就说这个点作等速运动.当圆弧上或一条直线上的一点,以等速运动通过两段线段,用来通过其中一条线段的时间与用来通过另一线段的时间之比,就等于这两条线段的长度之比"[14].用现代的术语来讲,奥托利可斯是说,一个点的速度是均匀的,如果它在相同的时间内通过相同的距离,并且,更进一步讲,如果这个点在时间 t_1 内通过距离 s_1,在时间 t_2 内通过距离 s_2,则有 $s_1 : s_2 = t_1 : t_2$.前几段中的一些命题,以及第三章中对割圆曲线的讨论都是源于这个定义.事实上,阿基米德在他的《论螺线》一文的一开始就对此作了非常详细的讨论,因为螺线本身就是定义为一个点

沿一线段作匀速运动,而这线段本身又同时绕着它的一个端点匀速旋转时,这个动点的轨迹.

希腊人肯定观察到了自由下落的物体并不是匀速的,因而他们意识到了加速度的概念.不过现存的明确谈论加速运动的文献很少,其中之一是来自公元 6 世纪对已失传的物理学家史特拉托(Strato)(公元前 3 世纪)的《论运动》一文的评注.史特拉托首先断言,自由落体"完成最后一段运动的时间最短",并且进一步指出它"依次通过相同距离越来越快"[15].换言之,加速运动意味着,依次通过相同距离所花的时间越来越少,因而其速度在不断增加.然而,从这些片言只语中并不清楚,史特拉托是否是指自由落体的速度与已下落的距离成正比.但在公元第 3 世纪,一位亚里士多德的评注者却说过"向下运动的物体随已落下的距离成比例地越来越快"[16].

虽然希腊人熟悉运动学的基本概念,但并没有任何迹象表明他们曾用其来进行数值计算,这与他们在天文学里的做法完全不同.另一方面,海伦的《量度》(Metrica) 一书却是实用测量手册方面的一本代表作,一本使读者能学会如何来测量各种图形的面积和体积的书.书中海伦讲述了如何得出数值结果来,即使其中包含了"无理的" 量.海伦有时也会给出证明,但他的目的主要是计算,而不是证明.这本著作在某种意义上让人想起中国及巴比伦的一些教本,但海伦经常引用像阿基米德和欧多克斯这样一些人的结果来为他的法则进行辩护.

《量度》的第一卷给出了用来计算平面图形的面积和立体图形的表面积的方法.在讨论了像长方形和等腰三角形这样一些简单情况之后,海伦接着讨论三边长给定后不规则三角形的面积求法.他给出了两个方法.第一个方法是以《原本》II - 12,13 题为基础的:给定三角形 ABC,作 BC(或其延长线) 的垂线 AD,再应用前面引用过的定理可证有 $c^2 = a^2 + b^2 \mp 2a \cdot CD$(图 4.30).由此可推知 CD,从而 $AD = h$ 也可得知.因此其面积$\frac{1}{2}ah$ 也可以算出了.

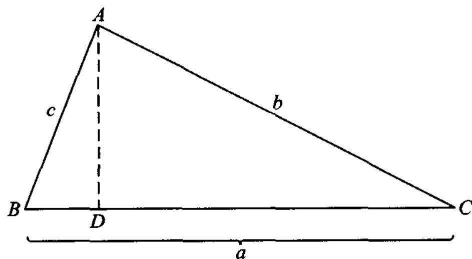


图 4.30 海伦求三角形面积的方法.

第二个方法就是我们今天所知的海伦公式.即,如果令 $s = \frac{1}{2}(a + b + c)$,则其面积等于 $\sqrt{s(s-a)(s-b)(s-c)}$.海伦是这样讲的,"设三角形的三边长分别为 7,8 和 9.将 7,8,和 9 加在一起,结果为 24.取一半,得 12.从 12 中减去 7 剩 5.从 12 中减去 8 剩 4;再从 12 中减去 9 剩 3.将 12 乘 5,结果为 60.再将此乘 4,结果为 240,再将此乘 3,结果为 720.取它的平方根,就得到这个三角形的面积"[17].

海伦在书中对这个面积公式给出了正确的几何证明.这个公式及其证明很可能是来源于阿基米德,这个公式在古希腊时代是很不寻常的,因为其中有四个长度的乘积,这完全是个"非几何" 的概念.海伦对这种似乎是有失常规的情况未作特别的说明,因此,这个结果很可能早在其它文献中已经出现过了,尽管在《原本》中只有两个或三个长度可以相乘,目的是得出长方形(的面积) 或长方体(的体积),但是对在本章中所讨论到的希腊数学某些方面的实际需要促使某些数学家把长度看成是"数",既然这样,当然就可以相乘.这个新观念自然违背了亚里士多德关于如何来理解数学的基本哲学信念.不过,这的确再次表明,希腊数学在其"几何外衣" 的后面又有了许多新的发展.

在这一节中海伦接着讲述了如何来计算所必须的平方根:

> 因为 720 没有有理的平方根,因此我来对这个根作近似计算,方法如下:由于最接近 720 的平方数是 729,它的平方根是 27.将 720 除以 27,得 $26\frac{2}{3}$;加上 27 是 $53\frac{2}{3}$;取其中一

半,结果得 $26\frac{5}{6}$. 因此 720 的平方根与 $26\frac{5}{6}$ 很接近. 由于 $26\frac{5}{6}$ 与自身相乘的乘积是 720 $\frac{1}{36}$,因而其差值是$\frac{1}{36}$. 如果想使差值小于$\frac{1}{36}$,我们就要用已求得的 720 $\frac{1}{36}$ 来替换 729,再用同样的方法,我们就能得到一个新的近似值,使得这时的差值远小于$\frac{1}{36}$.[18]

这个求平方根的方法是应用数学的又一个例子,有意思的是,它与塞翁所描述的托勒密方法有很大的不同,这也许是由于海伦的方法是用十进制来计算的,而托勒密的方法是用在六十进制的天文学计算中的.

《量度》一书还包括了边长为 a 的正 n 边形的面积 A_n 的一个公式,这里 n 的范围是从 3 到 12. 例如,海伦证明了 $A_3 \approx \frac{13}{30}a^2, A_5 \approx \frac{5}{3}a^2$,以及 $A_7 \approx \frac{43}{12}a^2$. 在每种情况下,他利用了各种平方根的近似值,这些平方根出现在几何推导中. 就是在他推导正九边形的面积时,海伦借助了"弦表". 从表中他发现了:40° 中心角所对的弦长等于该圆直径的三分之一. 因此 $AC^2 = 9AB^2, BC^2 = 8AB^2$(图 4.31),从而有 $A_9 = 9\Delta ABD = \frac{9}{2}\Delta ABC = \frac{9}{4}BC \cdot AB = \frac{9}{4}\sqrt{8}a^2 \approx \frac{9}{4} \cdot \frac{17}{6}a^2 = \frac{51}{8}a^2$.

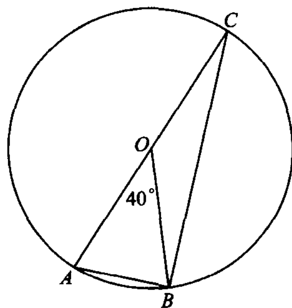


图 4.31 计算正九边形的面积.

为了求圆的面积,海伦用了阿基米德的$\frac{22}{7}$来作 π 的值,从而得出圆面积$\frac{11}{14}d^2$,这里 d 为圆的直径. 接着,他引用了"古人的"一个讲弓形面积的公式:$A = \frac{1}{2}\cdot(b+h)\cdot h$,这里 b 为底长,h 为弓形的高. 他还说,更精确的值应该再加一项$\frac{1}{14}\left(\frac{b}{2}\right)^2$. 取 $\pi = \frac{22}{7}$ 时这个新公式对半圆的确是精确的,但对其他弓形则只是近似的. 海伦甚至注意到这个公式只有当 $b < 3h$ 时才有比较合理的精度.

在《量度》第 Ⅱ 卷的前言中,海伦指出了长方体的体积为表示其长、宽、高的数的乘积,因为长方体可以分割成这么多个的单位立方体. 但是接着他讲述了更普遍的结论:在一个立体的图形中,如果它的任一平行于底面的截面都相等,而所有截面的中心都在一条通过底面中心的直线上,直线与底面或许是垂直或许是斜交,那么这图形的体积等于底面积与其在垂直方向的高度的乘积. 海伦既没能证明这法则,前面对长方体体积求法的说明在这里也用不上,因为一般来说,这个立体图形并不能分割成整数个单位立方体. 海伦大概已经认识到这种证明可通过不可分量的论证做到. 如果取一个长方体,它的底面积等于上述立体图形的底面积,它的高度也等于上述立体图形的高度. 那么由于每个立体是由它的"不可分的"平行截面"构成",而且由于这两立体的截面又相等,由此可知这两个立体的体积相等. 因为长方体的体积等于底面积与高度的乘积,因此对上面所说的立体图形这个结论也成立. 正如我们在前面几章中指出过的,这种不可分量的论证方法似乎在几个世纪前就在希腊数学里出现了,尽管这方法从来没有正式的地位,因此也从来没有"发表"过.

在卷 Ⅱ 的其余部分,海伦给出了计算许多其他立体图形体积的公式. 有些他只是简单地引述前人的结果,有些他则给出了一些初步的论证,虽然还不能算是正式的证明. 就在这些结果中他给出了一个环的体积的公式($2\pi^2ca^2$),其中 a 为圆形截面的面积,c 为截面中心到环中心的距离;他还给出了正八面体体积公式为$\left(\frac{1}{3}\sqrt{2}a^3\right)$,这里 a 为边长.

4.3.2 度量之书(The Mishnat ha-Middot)

在公元初的几个世纪里,海伦的教本在希腊 - 罗马世界广泛流传.同时,一些其他类似的著作也出现了,这些著作甚至更加远离欧几里得的模式.一个特别引起我们兴趣的例子是一本希伯来人的著作《度量之书》,其确切的年代和作者都不得而知,可是,人们普遍认为它是由第 2 世纪中期的拉比 · 尼希米亚(Rabbi Nehemiah)所著,可能是给巴勒斯坦的犹太农夫和工匠们用的.和海伦的《量度》一样,它也是一本有关计算面积和体积规则的汇编.但是,与《量度》不同的是,它里面没有证明.

现从其中摘录几小段,从中便可见它的风格[19].

> Ⅱ - 3　圆的面积如何算呢?…… 将直径自乘,再从中减去它的七分之一及十四分之一,所得结果便是圆的面积,例如,如果直径长为 7,自乘得 49,它的 $\frac{1}{7}$ 与十四分之一的和是 $10\frac{1}{2}$,因而面积为 $38\frac{1}{2}$.
>
> Ⅱ - 4　弓形的面积怎样算呢?…… 将矢长 * 与弦长相加,再将相加得到的结果乘以矢长的一半,将所得数放在一边;然后 …… 取弦长的一半,自乘后除以 14,将这样算得的数与上面算出放在一边的数相加,其结果便是弓形面积.

换句话说,圆面积公式为 $A = d^2 - \frac{1}{7}d^2 - \frac{1}{14}d^2 = \frac{11}{14}d^2$;而弓形面积公式为 $A = (b + h) \cdot \frac{b}{2} + \frac{1}{14}\left(\frac{b}{2}\right)^2$.这两个公式和海伦的公式是一致的.

> Ⅱ - 10,11　如何去计算(一平截头棱锥体的体积)呢?例如,取一底面为正方形的柱体,它的底面为 4 肘尺见方[它的高度为 10 肘尺],随着高度的增加(截面)逐渐减小,到了上端为 2 肘尺见方的正方形 …….你可以得出以下数据:当柱体升到整个棱锥一半高处时,其上、下底边长之比为 2 比 4,因此你将发现整个柱体直至顶点的高度为 20 肘尺,而至截头开始处的高度为 10 肘尺.

拉比 · 尼希米亚用标准的锥体的体积公式完成了他对高为 20 肘尺的整个棱锥体的体积的计算,结果得 $\frac{1}{3} \cdot 4^2 \cdot 20 = 106\frac{2}{3}$,而被截去的棱锥体的体积为 $13\frac{1}{3}$.因此,平截头棱锥体的体积为 $106\frac{2}{3} - 13\frac{1}{2} = 93\frac{1}{3}$. 海伦显然没有用这种方法,而是用了下面的公式

$$V = h\left[\left\{\frac{1}{2}(a + b)\right\}^2 + \frac{1}{3}\left\{\frac{1}{2}(a - b)\right\}^2\right],$$

式中 a, b 为下底和上底各边的长,而 h 为高.

> Ⅲ - 4　对于四边相等,但角度不等的四边形.面积该怎样算呢?例如,每条边长为 5,两个角小,两个角大,两对角线相互平分,长度分别为 6 和 8 的这样一个四边形.如果要计算其面积.则将一对角线的长度与另一对角线的一半相乘.得 24 肘尺即为所求.

这样一来,对角线长度分别为 d_1 和 d_2 的菱形的面积为 $A = \frac{1}{2}d_1 \cdot d_2$,在海伦的著作中也可以找到这个公式.类似地,尼希米亚用了海伦的公式来计算三角形的面积.

---

* 指弓形的高 —— 译者.

最后一个例子表明了拉比·尼希米亚曾力图克服科学与宗教信仰之间的冲突：

V－3　……如果你要知道绕(圆)一周的周长，则将直径乘以三又七分之一即可…….

V－4　圣经中这样写道“他铸造了一个圆形的铜盆，从边缘到边缘的距离为10肘尺”，然而其圆周长却是30肘尺，因为书中又写道：“而且一条长为30肘尺的线段正好把它包围起来”. 这里的“一条长为30肘尺的线段”是什么意思呢?尼希米亚解释说：“世人所说的圆周长等于直径的三又七分之一，但应减去其中的七分之一作为铜盆边缘的厚度，那么正好剩下30肘尺包围一周”.

我们在第1章中已经指出过，那句希伯来圣经中的引语是出自诸王本纪，7：23. 尼希米亚对这一圣经与“世人”之间的冲突提出的一个解决办法就是指出10肘尺的直径中一定是包括了铜盆的壁，而30肘尺的周长，则是除去盆壁后的内圆周长. 不幸的是，并不是总能这样简单地来解决宗教信仰与数学或其它科学研究之间的冲突.

习　题

选自托勒密的《大成》

1. 已知 $\mathrm{crd}(60°)=R=60$，利用希帕科斯的半角公式来计算 $\mathrm{crd}(30°)$，$\mathrm{crd}(15°)$ 及 $\mathrm{crd}\left(7\frac{1}{2}°\right)$，同时利用 $\mathrm{crd}(180°-\alpha)$ 的公式来计算 $\mathrm{crd}(120°)$，$\mathrm{crd}(150°)$，$\mathrm{crd}(165°)$ 及 $\mathrm{crd}\left(172\frac{1}{2}°\right)$.
2. 利用塞翁的方法计算 $\sqrt{4500}$ 到六十进制的小数点后第二位. 答案是67;4,55.
3. 利用托勒密关于圆内接四边形的定理来证明两角和公式
$$120\mathrm{crd}(180-(\alpha+\beta))=\mathrm{crd}(180-\alpha)\mathrm{crd}(180-\beta)-\mathrm{crd}\,\alpha\,\mathrm{crd}\,\beta.$$
4. 利用托勒密的差角公式来来计算 $\mathrm{crd}(12°)$，再用半角公式计算 $\mathrm{crd}(6°)$，$\mathrm{crd}(3°)$，$\mathrm{crd}\left(1\frac{1}{2}°\right)$ 及 $\mathrm{crd}\left(\frac{3}{4}°\right)$，将算出的结果与托勒密的结果相比较.
5. 写一个计算机程序来做习题4中的计算(用六十进制).
6. 将希帕科斯对半角公式的推导与阿基米德在《圆的测量》一书中推导引理2的方法相互比较一下.
7. 试证：$\mathrm{crd}\,\beta:\mathrm{crd}\,\alpha<\beta:\alpha$，或等价地 $\frac{\sin\beta}{\sin\alpha}<\frac{\beta}{\alpha}$，这里 $0<\alpha<\beta$.
8. 用类似于托勒密在求太阳方向例子中所用的方法，对 γ 为锐角情形导出余弦定律.
9. 用托勒密的方法，分别计算在纬度为40°和 $23\frac{1}{2}°$ 处，一长度为40的柱子在春分点午时日影的长度.
10. 说明为什么黄道与赤道的交角 ε 等于太阳在夏至和冬至午时的高线的角距离的一半(见图4.32).

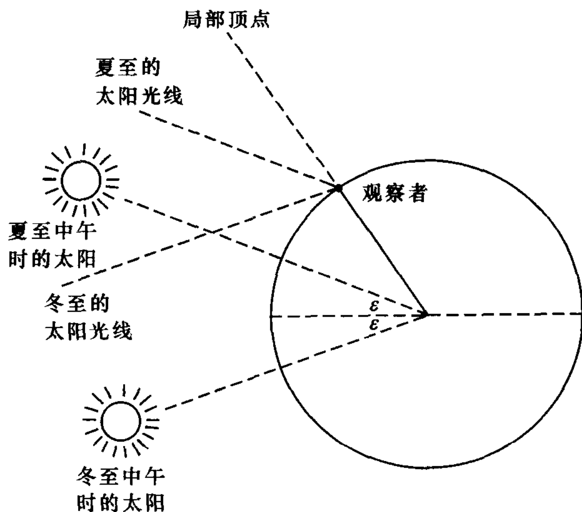


图4.32　计算黄道的倾角.

11. 计算一位于纬度 $23\frac{1}{2}°$ 和 36° 处一根长为 60 的立柱在夏至和冬至时日影的长度(在图4.33中,G代表太阳在夏至午时的位置,B为春分午时的位置,L为在冬至午时的位置,而且根据习题10,弧 GB = 弧 $BL = 23\frac{1}{2}°$).
12. 计算当太阳经度分别为 90°(夏至),120° 和 45° 时,它的赤纬和赤经值.根据对称性计算经度分别为 270°,240° 和 315° 时,太阳的赤纬值.
13. 写一计算机程序,用来计算经度为任意值 λ 和地理纬度为 φ 的地方太阳升起的时间 $\rho(\lambda,\varphi)$.

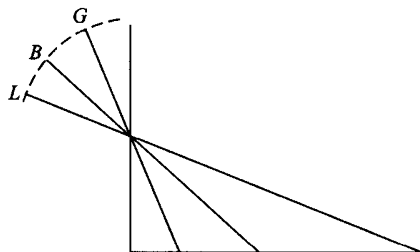


图 4.33　日影的长度.

14. 计算 $\phi = 45°, \lambda = 60°$ 和 90° 这两地的日出时间 $\rho(\lambda,\phi)$.
15. 计算在纬度为 36°,$\lambda = 60°$ 处白昼的长度,并计算当地的日出和日落时间.同样,在纬度为 45°,$\lambda = 60°$ 的地方计算其白昼长度和当地的日出和日落时间.
16. 设已知在某一地点白昼最大的长度为 15 小时,计算该处的纬度及夏至和冬至时日出的位置.
17. 公式 $\sin\sigma = \tan\delta \cdot \tan\phi$ 仅当等式右边值小于或等于1时才有意义.由于角 δ 的最大值是 $23\frac{1}{2}°$,证明当 $\phi > 66\frac{1}{2}°$ 时,右边的值总是大于 1.试在此情况下用白昼长来解释这个公式.
18. 在纬度为 45° 处,分别计算 $\lambda = 45°$ 和 90° 时,太阳离天顶的角距离.
19. 在地理纬度 20° 处,大约在哪一天,午时太阳正好直接在头顶上?
20. 分别对纬度 45°,36° 和 20° 计算最北的日出点.在纬度为 75° 的地方,大约在什么日子会开始出现"午夜太阳"?

选自海伦的著作

21. 说明应如何利用相似三角形来计算两个达不到的点之间的距离(例如,假设两个点都在河的另一面).
22. 用海伦的方法计算边长分别为 4,7 和 10 的三角形的面积.
23. 在海伦的边长为 a 的等边三角形的面积公式 $A_3 = \frac{13}{30}a^2$ 中,他对$\sqrt{3}$ 用了什么样的近似值.用他的求平方根的算法导出这个近似值.
24. 试导出边长为 a 的正五边形的面积公式 A_5(用平面几何).讨论所导出的公式与海伦公式 $A_5 \approx \frac{5}{3}a^2$ 之间的差别.
25. 海伦在推导边长为 a 的正七边形面积公式 $A_7 = \frac{43}{12}a^2$ 时,利用了假设 $a = \frac{7}{8}r$,这里 r 为外接圆的半径.利用这个近似导出海伦的结果.这里必须用什么样的平方根近似?
26. 导出$\sqrt{8}$ 的近似值为 17/6,并完成 A_9 的海伦公式的证明.
27. 利用三角学导出边长为 a 的正 n 边形的一般面积公式 A_n.
28. 对边长为 a 的正八面体,导出海伦的体积公式$\frac{1}{3}\sqrt{2}a^3$.

讨论题

29. 按照托勒密讲解的顺序,勾勒出三角学教程的纲要.也就是说,先导出一些主要的公式作为编制正弦表的工具.讨论这种方法与现在标准教材中所采用的方法相比的优缺点.
30. 托勒密一定知道某种用圆锥曲线来三等分一个角的方法.这种方法使他能够用已知 $1\frac{1}{2}°$ 的对弦值来得出 $\frac{1}{2}°$ 的对弦值.为什么托勒密不把这种方法看成是用"几何方法"来作图的方法?是否可以用这种作图法来计算出 $\frac{1}{2}°$ 对弦的数值?

31. 讨论是否可以按照托勒密的思路,在三角学的课程中加入部份球面三角学的内容.
32. 设计一堂课,用一些基本的球面三角关系来计算一些简单的天文现象.
33. 有哪些观察结果使古希腊人相信,天球半径很大,以致于地球相对于天球来说可以看成是一个点.
34. 说明有哪些证据使你相信:(a) 地球沿自身的轴每天旋转一周,(b) 地球围绕太阳每年旋转一周.这些证据说服了古希腊人吗?你如何来驳斥托勒密所提示的认为地球是不动的理由?
35. 在一本天文学的著作中找出"时间的方程",并讨论一下,为什么用书中的方法计算日出与日落的时间会有好几分钟的误差.
36. 阅读一下 1990 年刊登在《Mathematics Magazine》上的一篇文章,从中可以了解到,虽然夏至与冬至的确是一年中最长和最短的两天,但是日出,日落时间的极值却不是在这两天出现[20].写一篇简短的报告来解释这一现象.

文献和注解

在第 2 章中提到过托马斯·希思(Thomas Heath)的 *A History of Greek Mathematics* 以及 B. L. 范·德·瓦尔登(van der Waerden)的 *Science Awakening*,这二书中均包含有与本章有关的材料.托勒密等人著作的节选可以从 Ivor Thomas 的 *Selections Illustrating the History of Greek Mathematics* 中找到.而论述从巴比伦时代到公元第6世纪期间数学在天文学中的应用的标准文献则是奥托·诺伊格鲍尔(Otto Neugebauer)所著的 *A History of Ancient Mathematics to astronomy*(New York:Springer,1975).该书详细地叙述了托勒密和其他天文学家们在研究他们各自的天文系统时所采用的一些数学技巧.有关托勒密的《大成》的英译本.如今能找到的最好的版本是由 Gerald T·Toomer 翻译的 *Ptolemy's Almagest*(New York,Springer,1984).由 Gateaby Taliaferro 翻译的较早的译本则可在 *Britanica Great Books* 中找到.

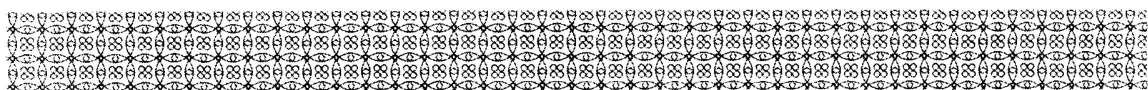
1. Simplicius 评注的亚里士多德的 *On the Heavens*,转引自 Pierre Duhem 的 *To Save the phenomena* 一书(Chicago:University of Chicago Press,1969).p.5. *Duhem* 的著作详细地讲述了古代希腊人是如何勉力去"拯救表观现象"的.
2. 托勒密,上述著作,Ⅰ,7.
3. Thomas Heath, *The Works of Archimedes*(New York:Dover,1953),p.222.
4. Plutarch 的"论月球的表面",见 Thomas, *Selections Illustrating*, Ⅱ.p.5.
5. 这一讨论是摘自托马斯·库恩(Thomas Kuhn)的 *The Copernican Revolution*(Cambridge:Harvard University Press,1957),p.58.该书提供了有关古希腊时代天文学的本质的极好的背景阅读材料,更近期的,特别是为本科生设计的著作则是由 Michael J. Crowe所著之 *Theories of the World from Antiquity to the Copernican Revolution*(New York:Dover Publications,1990).而对星球运动的详尽的数学描述,包括一些图解,则可从诺依格鲍尔的《古代天文数学史》(pp.677 – 685)中找到.
6. Theon, *Commentary on Ptolemy's Syntaxis*.采自 Thomas, *Selections Illustrating*, Ⅰ.p.61.
7. 托勒密的《大成》.
8. 下面的讨论是采自 J.L Bergren 的论文"Mathematical Methods in Ancient Science:Astronomy",该文收在 Ⅰ.Grattan Guinness 主编的 *History in Mathematics Education*,(Paris:Belin,1987)一书中.
9. 关于梅内劳斯定理的证明,托勒密的《大成》Ⅰ.13,或诺依格鲍尔的 *History of Ancient*,pp.27 – 28.
10. Olaf Pedersen, *A Survey of the Almagest*(Odense:University Press,1974),p.93.本书是对托勒密著作译本的极好的辅助材料.它对托勒密书中所有数学和天文学方面的内容都提供了背景材料和注释.
11. 托勒密,《大成》,Ⅸ,2
12. 欧几里得, *L'Optique et la Catoptrique*,由 Paul Ver Eecke 译成法文(巴黎,Descleé de Brouwer,1938),p.13.
13. Thomas, *Selections Illustrating*, Ⅱ.p.497.
14. 引出 Marshall Clagett, *The Science of Mechanics in the Middle Ages*(Madison:University of Wisconsin Press,1961),p.165.虽然这本书主要是讨论中世纪的力学的,但在大多数的章节的开头部分都有对希腊人的工作的一个总结.
15. *Simplicius*, *Commentrary on Aristotle's Physics*,收在 Morris Cohen 与 J.E.Drabkin 编的 *A Source Book in Greek*

Science(Cambridge:Harvard University Press,1948)一书中,p.211.该书是一本极好的有关希腊数学,天文,物理和其它科学方面的原始资料汇集.

16. 引自 Clagett 的 *Greek Science in Antiquity*(New York:Collier,1963),p.92.Clagett在本书中对希腊科学的各个方面,从它的一开始,经过对拉丁科学的影响,一直到中世纪的早期,都作了简洁的描述.
17. Thomas, *Selections Illustrating*, Ⅱ.p.471.
18. 同上 Ⅱ.p.471.
19. 选用 Solomon Gandz 的 *Studies in Hebrew Astronomy and Mathematics* 一书中的《度量之书》(*Mishnat ha-Middot*),KTAV Publishing House,1970,经许可重印.Gandz在引言中,对为何把这本书放在第2世纪提出了自己的理由.更新近的研究有 Gad B Sarfatti 的"The Mathematical Terminoloey of the *Mishnat ha-Middot*", *Leshonenu* 23,156 - 171,特别从所用的语言上来分析,对 Gandz 的工作提出了疑问,并提议说该著作是在9世纪以后才形成文字的.然而,这本书与海伦的《量度》的惊人相似这一点,令人相信作者是熟悉海伦这本书的,很可能是将它改编成了一本适合第2世纪巴勒斯坦国情的书.
20. Stan Wagon,"Why December 21 is the Longest Day of the Year," *Mathematics Magazine* 63(1990),307 - 311.

天文学与实用数学概览

公元前408—335年	欧多克斯(Eudoxus)	宇宙二元模型
约公元前300年	奥托利可斯(Autolycus)	球面学的教材
约公元前300年	欧几里得(Euclid)	球面学和光学教材
公元前310—230年	阿利斯塔克(Aristarchus)	到太阳和月球的距离
公元前250—175年	阿波罗尼乌斯(Apollonius)	偏心圆和本轮体系
公元前190—120年	希帕科斯(Hipparchus)	三角学的开始
公元第1世纪	海伦(Heron)	测量学、力学、光学
约公元100年	梅内劳斯(Menelaus)	球面三角学
约公元100—178年	托勒密(Ptolemy)	《大成》
公元第2世纪	拉比·尼希米亚(Rabbi Nehemiah)	《度量之书》
公元330—405年	塞翁(Theon)	《大成》评注



第5章 希腊数学的晚期

> 这块墓地里躺着丢番图……[而且它]科学地告诉我们他的生命的历程.上帝赐给他的生命的六分之一做一个男孩,然后,加上他生命的十二分之一,他的两颊开始生出细软的胡须.再过了生命的七分之一,上帝为他点燃了婚姻的烛光,又在他婚后第五年时,赐给他一个儿子,天哪!这个晚生的可怜的孩子:在达到了他父亲生命的一半的时候,残酷的命运之神把他带走.在他用这数的科学安慰自己、悲伤地度过4年之后,他结束了自己的生命.
>
> (《希腊作品选》(约公元50年)卷XIV 之警句126)[1]

埃及,公元415年3月:"(于是)在基督徒中流传着这样一个谣言,说塞翁的女儿[希帕蒂娅]是阻止省长[奥热斯特斯(Orestes)]和主教[西里尔(Cyril)]和解的惟一障碍.在四旬斋期间一个不幸的日子里,希帕蒂亚被人从车子里拉了出来,剥光衣服,拖到教堂里,被诵经师彼得(Peter)和一帮野蛮残忍的狂热分子惨无人道地屠杀…….杀害希帕蒂亚无论是在人格还是在宗教上都永远将亚历山大的西里尔钉上了历史的耻辱柱"[2].

在埃及-希腊托勒玫王国的统治下,到处都兴建博物馆和图书馆,尽管托勒玫王国的统治于公元前31年在罗马势力的攻击下垮台,但亚历山大的理智-科学的传统继续存在了许多世纪.在第4章中我们讨论了三位杰出的埃及数学家的工作,他们是海伦、梅内劳斯和托勒密,还有一位来自罗马属地的朱迪亚的拉比,他们都活跃于罗马人统治下的埃及.在公元初的几个世纪里,还有其他一些数学家,他们的著作的影响一直延伸到文艺复兴时期.本章将讨论其中四位(补遗5.1).

补遗 5.1

罗马的数学

长期以来都认为不存在"罗马数学".自然,有一些有创造性的数学家在罗马帝国工作,主要是在亚历山大,但是他们仍然是属于希腊传统的一部分.然而,文献中没有记载有任何数学家居住和工作在帝国的中心或是用拉丁文来写作.伟大的演说家西塞罗(Cicero)甚至承认,罗马人对数学不感兴趣:"希腊人把几何学家捧得至高无上,因此在他们看来,没有什么能比数学造成更辉煌的进步.但我们却指出了这门技艺的局限,它只是在测量和计数上有用"[3].

实际上,罗马人还真有更多地应用数学的地方,不仅仅是"测量和计数"而已.正如维特鲁维乌斯(Vitruvius)在《论建筑》中写的"数学能为建筑提供许多资源,它教会人们使用直尺和指南针,而这又有利于通过用直角尺,水平尺和准线来设计安排.利用建筑光学,就可以保证每天光线能从天空的某些方向射进室内.利用算术,可以计算建筑费用;借助于几何法则,可以解决困难的对称问题,从而指明测量的方法.[4] 即使这样,建筑师们似乎只需要算术、几何和光学方面的一些零碎的知识就够了.所有必需的知识都可以从像海伦的著作中获取.实际上在维特鲁维乌斯著作中提到的数学概念是很少的,包括在建筑中用到的某些比例,二倍平方和立方问题,以及可以精确地作出直角的、由三根长为3,4,5的杆构成的直角尺.

对测量来说,需要什么样的数学呢?[5] 罗马的测量师设计了横穿辽阔帝国的道路和水渠,其中许多至今还存在.但是查看现存的一些当时的测量手册,发现罗马测量师只用了一些非常初等的数学.卢修斯·哥伦梅拉(Lucius Columella),一个公元1世纪上半纪的地主,曾经写道:一个和田地打交道的人必需会算面积.因此他给出了一些计算正方形、长方形、圆等的面积公式,这些我们已经在海伦的著作中见到过了.他还用了22/7来作为π的近似值,用了

$$A = \frac{1}{2}(h+b)h + \frac{1}{14}\left(\frac{b}{2}\right)^2$$

来计算底为b,高为h的弓形面积.由惠格莫斯·吉奥马弟卡斯(Hygmus Giomaticus)所编的测量手册中有一个确定正北方向的方法.在一块平地上划一个圆,在其中心放置一个日规的指针,指针要足够长,使得它的阴影有时能落在这个圆的外面.然后分别标记出上午及下午阴影穿越圆周时与圆周的交点,在这两点之间连一直线,然后作该直线段的垂直等分线,这条等分线就会指向正北和正南方(图5.1).

由马卡斯·琼尼乌斯·尼普修斯(Marcus Junius Nipsius)所著之手册中叙述了用全等三角形来测量一条河的宽度的方法(图5.2),要求的是距离BC.点A与BC在一条直线上,线AD与AC正交,G为AD的等分点.从点D作AD的垂线至点H,使HGC在一直线上.则BC等于$DH-AB$.可见,罗马的测量师们显然采用了甚至比海伦的方法更初等的方法.

看来用于测量、建筑或其他管理这个帝国所必需的事务中的数学,全都取自早先的发现,它们似乎已足以解决任何出现的问题,没有更多的需要.对那些在这个领域中满足智力好奇心的追求没有官方的鼓励,罗马帝国竟然在这样的情况下生存了500年,对这个世界的数学知识宝库没有任何贡献.

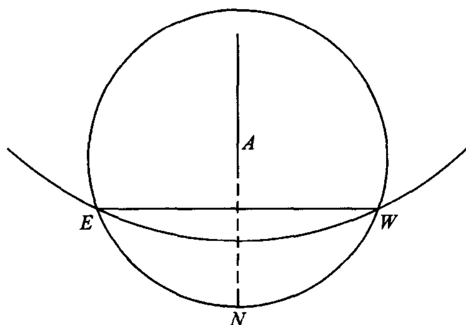


图 5.1 确定正北方.

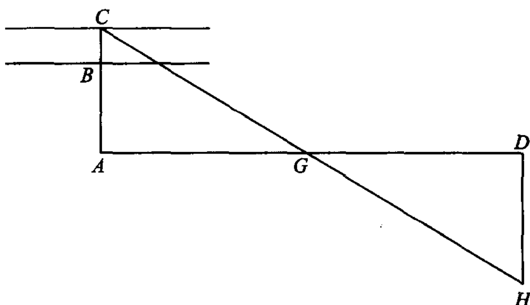


图 5.2 求跨越一条河的距离.

我们首先来谈尼可马科斯(Nicomachus)著作,他是希腊朱迪亚地方杰拉沙(Gerasa)镇上的人.他在公元第1世纪的晚期,在对毕达哥拉斯数的哲学理解的基础上写了一本《算术入门》的书.除了欧几里得《原本》中的第7-9卷外,这是现存惟一的古希腊数论著作.然而,还有另一本重要的著作,书名叫《算术》,是由亚历山大的丢番图(Diophantus)在第3世纪写就的,它注定要成为一本远比尼可马科斯的书重要得多的著作.尽管书名叫算术,它讲的却是代数,内容主要是将问题有机地组织汇集起来,这些问题翻译成今天的语言,就叫做不定方程,全都要求用有理数来解.和海伦的《量度》一样,《算术》一书的风格更像一本中国或巴比伦的习题汇编而不像经典的希腊几何著作.我们要谈的第三位数学家也是来自亚历山大,他就是4世纪初的几何学家帕普斯(Pappus).他所以出名不是由于他的原创性的研究工作,而是由于他对希腊数学的各个方面所作的注释,特别是在对几何分析中的希腊方法的讨论.本章最后将对希帕蒂娅(Hypatia)的工作作一简单的讨论,她是我们了解得最全面的第一位女数学家.她死于一群狂怒的暴徒之手,她的死标志着希腊数学传统在亚历山大的实际终结.

5.1 尼可马科斯和初等数论

我们对尼可马科斯的生平几乎一无所知,但是由于他的工作充满了毕达哥拉斯的思想,他很可能在亚历山大学习过,亚历山大是数学活动和新毕达哥拉斯哲学的中心.他的两本著作被保存下来,一本是《算术入门》;另一本是《调和比例入门》.从别的资料来源看,他似乎还写过几何入门和天文学入门,这样他就完成了讨论柏拉图的四门基本课程(所谓四艺)的一个系列.

尼可马科斯的《算术入门》一书可能是花了多年写就的,是解释毕达哥拉斯数的哲学的几本著作中惟一保存下来的一本.但是因为不存在来自毕达哥拉斯时代的教本,它就成了了解毕达哥拉斯数论中若干基本概念的史料来源,这在第2章中已经讨论过.可是,因为该书是在毕达哥拉斯之后600年写的,所以还必须结合它的时代来考虑,同时还必须将它与其它惟一还可以看到的数论著作——欧几里得《原本》中的第7-9卷——相比较.

尼可马科斯的这本书写成两卷,以一篇哲学的引言作为全书的开篇.和欧几里得一样,他仿照亚里士多德,区分了连续的"量"(magnitude)和不连续的"数"(multitude).和亚里士多德一样,他注意到后者通过不断地增加而成为无穷,而前者则通过不断分割而成为无穷.他进一步对四艺中的四

门学科进行划分,把算术与音乐同几何和天文学区分开来:算术和音乐处理的是离散的量,其中前者是绝对的,后者则是相对的;而几何与天文学处理的是连续的量,其中前者是静止的,后者则是运动的.在这四门学科中首先应该学习的是算术,"这不仅因为 …… 它作为某种普遍的模式比其它几门学科更早地存在于造物主上帝的心目中,造物主以它作为范例和蓝图,将所创造出来的宇宙万物安排得井然有序,并使它们达到各自的最佳终结,而且还因为它按本性来说就先于其他学科而生,即其他学科会因它的衰亡而衰亡,而它却不会与别的学科一起衰亡"[6].换言之,算术是其他三门学科所必需的.

尼可马科斯的《算术入门》一书第 1 卷大部分是讨论整数的分类和它们之间的关系.例如,作者把偶数分成三类:偶数乘偶数类(二的幂)、偶数乘奇数类(奇数的二倍数),以及奇数乘偶数类(所有其他的偶数).奇数再分为素数和合数.在我们看来,尼可马科斯在讨论这些分类和说明各种数的形成方面所用篇幅似乎有些过多.但我们要记住,他是为初学者写的一本入门书,而不是为数学家写的专著.

尼可马科斯讨论了用互易相减的欧几里得算法来求两个数的最大公约数,以及来确定两个数是否互为素数.他还讨论了完全数的问题,用的也是欧几里得构造法(《原本》IX,36),但和欧几里得不同的是,他实际算出了前四个完全数为 6,28,496 和 8128.然而,又是和欧几里得不一样,尼可马科斯没有提供证明,只是给出了例子.

第 1 卷的最后六章讨论命名不相等数的比率的一个精细的十重分类方案,这个方案很可能源于早期的音乐理论.这个方案在中世纪和文艺复兴时期的算术中是很常用的,而且有时能在欧几里得《原本》的早期印刷本中找到.在这个命名比率 $A:B$ 的分类方案中,在化简到最低项而写成 $a:b$ 后,当 $a = nb$ 时叫做"倍率";而当 $a = b + 1$ 时叫做"定超率"(Superparticular),当 $a = b + k(1 < k < b)$ 时叫做"分超率"(Superpartient).

然而,最使我们感兴趣的是该书的第 2 卷,因为在本卷中他讨论了平面数与立体数,虽然讨论得很详细,但同样没有证明.这方面的材料欧几里得根本没有提到.尼可马科斯不仅讨论了三角形数和正方形数(见第 2 章),还讨论了五边形、六边形和七边形数,而且说明了如何无限地扩展这个系列.例如,五边形数是 1,5,12,22,35,51,⋯(虽然尼可马科斯在此注意到了 1 仅仅是一个"潜在的"五边形的边长).这些数用第 2 章中的点记号画出来成为一些等边的五边形(图 5.3).从 5 开始,这个数列中的每一项均可由前一项加上一相关数列 4,7,10,⋯ 中的下一个数来得出.这样,$5 = 1 + 4$,$12 = 5 + 7$,$22 = 12 + 10$,等等.这和三角形数 1,3,6,10,⋯ 的情形完全相似,这个数列中的每一项由前面一项依次加上数列 2,3,4,⋯ 中的数形成.这也和正方形数列 1,4,9,16,⋯ 的情形相似,这个

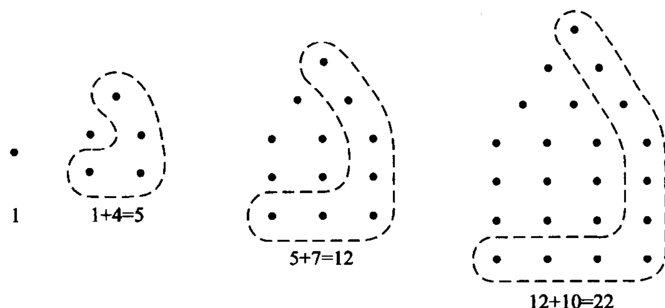


图 5.3　五边形数.

数列中的每一项由前面的一项依次加上数列 3,5,7,… 中的数构成.尼可马科斯继续这样类似地做下去,并且写出了上述每一种多边形数的前 10 个数.

尼可马科斯更进一步研究了立体数.以一给定 n 边形为底的棱锥体数(Pyramidal number),由前 n 个该种形状的多边形数相加形成.例如,以三角形为底的棱锥体数列为 $1,1+3=4,1+3+6=10,1+3+6+10=20,\cdots$,而以正方形为底的棱锥体数列为 $1,1+4=5,1+4+9=14,1+4+9+16=30,\cdots$.我们可以类似地构造出以任意多边形为底的棱锥体数列.

另一类立体数为立方数.尼可马科斯指出了(又未给证明):立方数是由奇数而不是由偶数形成.例如,第一个(潜在)立方数1,等于第一个奇数.第二个立方数8由接下去的两个奇数之和形成,第三个立方数 27 等于接下去的三个奇数之和.因此立方数与平方数紧密相关,后者也是由奇数相加构成.从而,尼可马科斯得出结论,这两个事实说明,奇数,而不是偶数,是"同一性(Sameness)"的根源.

该书最后一个论题是比例.尼可马科斯对比例一词的使用参考了欧几里得以前的术语,其含义与欧几里得《原本》卷 Ⅶ 的定义有所不同.在欧几里得那里,三个数,如果第一个数作为第二个数的倍数(或分数)值,与第二个数作为第三个数的倍数值相等,就说它们成比例.尼可马科斯指出,"古人"不仅考虑了这种类型的比例(他把这叫做几何的比例),还考虑了另外两种,即算术比例和调和比例.对尼可马科斯来说,三项成**算术比例**是指每一对相邻项之差相同.例如 3,7,11 成算术比例.这一比例有这样一个性质:它的两端的值之积等于平均值的平方减去它们之差的平方.在**几何**比例中("这是惟一能在这个词的严格意义上被称之为比例的"[7]),最大的项与次大项之比等于这个项与再次一项之比.例如,3,9,27 成几何比例.这个比例有一个性质,就是它们两端的值之积等于其中项的平方.在这方面尼可马科斯引述了欧几里得的结论,即在两个平方数之间只有一个平均项,而在两个立方数之间有两个平均项.

在三个项之间的第三种比例叫**调和比例**,在这种比例中最大项与最小项之比等于最大项与中项之差比中项与最小项之差.例如,3,4,6 这三个数成调和比,因为 $6:3=(6-4):(4-3)$,这种比例有这样一个性质,就是它的两端的值相加后乘中项值正好为两端值乘积的二倍.尼可马科斯对于"调和"这个名称作了一个可能的解释,就是6,4,3来自最基本的和声.比例 $6:4=3:2$ 给出第五度音阶,而 $4:3$ 给出第四度音阶,而 $6:3=(4:3)(3:2)=2:1$ 给出第八度音阶.今天,更普通的用法是将"算术的"、"几何的"和"调和的"这些名称用到均值上,而不是用到比例上.例如,7是3和11的算术平均值,9 是 3 和 27 的几何平均值,而 4 是 3 和 6 的调和平均值.

《算术入门》一书显然只不过是有关正整数的一些基本概念的导引.尽管它与欧几里得的《原本》有一些共同之处,然而它是在一个低得多的水平上写成的.其中没有任何证明,而只有一大堆例子.因此这本书适合在学校里的初学者使用.事实上它在古代得到了广泛的传播,在 9 世纪被译成了阿拉伯文,并被博伊修斯(Boethius)(约 480—524) 意译成拉丁文,于中世纪初期在整个欧洲被采用.由于这个原因,有些版本仍然存在.这本书是这样流行,在欧洲这一时期的大部分时间里,没有人学过比这更高深的著作,包括欧几里得《原本》,这就表明了数学学习的水准从希腊的全盛时期之后已经下降了.这些初等的数论性质在好几个世纪里都是算术课程的顶峰.

5.2 丢番图和希腊代数

对丢番图的生平我们知之甚少,除了在本章开始引用的墓志铭上的那些内容,我们只知道他生活在亚历山大.他的影响通过他的主要著作《算术》而绵及现代.丢番图在他的引言中告诉我们,

《算术》共分为十三卷.只有六卷在希腊保存下来了,另外有四卷最近发现了其阿拉伯文的版本,从其内容来看,它似乎是整部著作中的第4到第7卷,随后才是希腊文的后三卷[9].我们以下将把希腊文的六卷称为 Ⅰ – Ⅵ 卷,而把阿拉伯文的四卷分别称为 A, B, C 和 D.阿拉伯文的那几卷在风格上与希腊文的那几卷略有不同,其中对一个问题的解的每一步说明得更充分.因此,很有可能,这本阿拉伯的著作不是丢番图原著的译本,而是从希帕蒂娅在公元400年左右所写的《算术》的注释本翻译过来的.

在讨论《算术》中的具体问题之前,应该提出丢番图在方程求解方面作出的重大进步,这就是他引进了符号.埃及人和巴比伦人是用文字来表述方程.丢番图则使用缩写符号来表示方程中的各项(见补遗5.2).相对于希腊传统而言,一个明显的突破是:他处理了高于3次的幂.

补遗 5.2

丢番图使用的术语和符号

"所有数都是由若干个单位组成的……",这些数有:

平方数　它是由任何数自乘形成;这个任意的数就叫做这个平方数的边;

立方数　它由平方数与其边的乘积形成;

平方 – 平方数　它由平方数自乘形成;

平方 – 立方数　它由平方数和由相同的边表成的立方数相乘得出;

立方 – 立方数　它由立方数自乘形成.

通过对这些数加、减或乘,以及它们之间的比或与它们的边的比,大多数的算术问题就这样形成,再依照下述的方法,你就可以解决这些问题.

现在这些数已经被公认为算术科学的基础,它们的简称已经给出:[未知量的]平方叫做 dynamis,用带有上标 Y 的 Δ,即 Δ^Y 来表示;[未知量的]立方叫做 Kubos,它的符号就是 K^Y;平方的自乘叫做 dynamo-dynamis,它符号就是 $\Delta^Y\Delta$;平方乘以由相同的根形成的立方叫做 dynamo-kubos,记成 ΔK^Y;立方的自乘叫做 Kubo-Kubos,记成 K^YK.

没有上述特性,只含有不确定的单位个数的数叫 arithmos(数),记成ς.还有一个用来表示确定数中的不变成份即单位(unit)的记号:M 上带有指标0,即 $\mathring{M}$.

注意丢番图所用的全部符号都是缩写,包括最后两个:ς为 $\alpha\rho\iota\theta\mu o$ς(arithmos或数)头两个字母的缩写,而 $\mathring{M}$ 代表 $\mu o\nu\alpha$ς(monas或单位).例如,原稿中有这样的式子 $\Delta^Y\gamma$ ς$\iota\beta\mathring{M}\theta$,它表示3个平方、12个数和9个单位,让我们写出来就是 $3x^2+12x+9$(记住希腊人用字母符号来代表数字,例如,$\gamma=3$, $\iota\beta=12$ 以及 $\theta=9$).丢番图还进一步使用记号 χ 来表示倒数,例如,$\Delta^{Y\chi}$ 表示 $1/x^2$.还有,符号⋔用来表示"负",这可能是来自 $\lambda\epsilon\iota\psi\iota$ς(Lepsis即负)的缩写.例如,$K^Y\alpha$ ςγ ⋔$\Delta^Y\gamma\mathring{M}\alpha$ 表示 x^3-3x^2+3x-1(负项总是归并在一起,这样单个⋔就足以表明它后面的所有项都是负的).不过,在讨论丢番图问题时,我们还是会用现代的记法.

丢番图也认识到了乘以负号的规则:"负号乘以负号得正号,而负号乘以正号为负号"[10].自然,丢番图在这里谈的不是负数,对他来说还不存在这样的概念.他只不过是叙述乘以含有减法的代数表达式时所必需的规则.丢番图假定具有正负项的加、减法则均为已知.在他快结束引言时写道:

> 对那些开始学习本书的人来说,最好是已经有了将各种类型的项进行相加、相减和相乘的实践,他必须能将具有不同系数的正的或负的项加到别的项上去,而这些项本身或者是正的或者部分是正的,部分是负的.他也还必须会从正项与负项的组合中减去别的项,

而这些项可能是正的,也可能是部分正、部分负.

接着,就解方程的基本法则简明叙述如下:

如果一个问题引出了一个方程,其中某些项等于一些同种类型但不同系数的项,这就有必要在等式两边从同类项中减去同类项,直至得到一项等于一项的形式.如果碰巧有这样的情况,即在等式的一边或两边有负的项,就有必要在两边都加上负的项,直至两边都是正的项,然后再从同类项中减去同类项,直到两边都只剩下一项.这应该是表达命题的假设时就确定的目标,就是说,如果可能的话,要将方程一直化简到只剩一项等于一项.不过,后面我会向你们说明,在剩下两项等于一项的情况下如何来解这样的问题.

换言之,丢番图解方程的普遍方法就是引导到一个形如 $ax^n = bx^m$ 的方程,其中 m 和 n 不大于 2,这至少在该书的前三卷是如此.另一方面,他又的确知道如何去解,譬如说,形如 $ax^2 + c = bx$ 这样的方程.

5.2.1　线性方程和二次方程

丢番图的大多数问题都是不确定的,即它能写成一组 k 个方程,而未知量的个数却多于 k.通常是有无限多个解.对这种问题,一般来说,丢番图只给出一个解,但我们可以很容易扩展他的方法得出其它解.对于确定性的问题,一旦某些量明确地给定了,则只会有一个解.下面将叙述这样两类问题的例子.

问题 Ⅰ-1　将一给定的数分成有给定差值的两个数.

丢番图在给定数为 100,给定差为 40 的情况下,给出了这个问题的解.设 x 为这两个数中较小的那一个,则 $2x + 40 = 100$,因而 $x = 30$,而所要求的数为 30 和 70.一旦"给定"数被指定,这就成了一个确定性的问题,但是丢番图的方法对任意的一对数都适用.如果 a 为给定数,$b < a$ 为给定差,则方程就是 $2x + b = a$,而所要求的数就是$\frac{1}{2}(a - b)$和$\frac{1}{2}(a + b)$.

问题 Ⅰ-5　将一给定数分成两个数,使得每一个数的给定分数(不一定要相等)相加之后得出一给定数.

用现代的记号来表示,这个问题就是这样的:已给 $a,b,r,s(r < s)$,要求我们找出两个数 u,v,使得有 $u + v = a$,$\frac{1}{r}u + \frac{1}{s}v = b$(丢番图在这里认为分数就是单位分数,他经常是这样做的).丢番图指出,要想这个问题有解,必须有$\frac{1}{s}a < b < \frac{1}{r}a$.接着他提出了在 $a = 100, b = 30, r = 3$ 和 $s = 5$ 的情况下的解:设(100 的)第二部分为 $5x$,因此第一部分为 $3(30 - x)$.从而有 $90 + 2x = 100$ 和 $x = 5$.于是要求的两部分为 75 和 25.

和问题 Ⅰ-1 一样,一旦"给定"的数被指定了,这个问题就变成一个确定性的问题,因而这个方法对任何满足所要求的条件的"给定数"都适用,丢番图取未知量为第二部分的 1/5.这就使得他在做后面的计算时避免了分数,因为第一部分的 1/3 应等于 $30 - x$,从而第一部分就是 $3(30 - x)$.解的其余过程是很清楚的.为了检验其方法的一般性,令 sx 表 a 的第二部分,$r(b - x)$ 为第一部分,方程化为 $sx + r(b - x) = a$ 或 $br + (s - r)x = a$.那么 $x = \frac{a - br}{s - r}$ 就是一个一般解.由于 x 必须为正,所以 $a - br > 0$ 或 $b < \frac{a}{r}$,这是丢番图的必要条件的前半部分.由 $sx < a$ 或 $s\left(\frac{a - br}{s - r}\right) < a$ 可以推得$\frac{1}{s}a < b$ 或 $a < sb$,这就是必要条件的第二部分.在这个特殊的问题中,给定值是这样来挑选

的,使它能保证答案为整数.在第1卷中多数问题都是这样做的.但在其他几卷中,对解加上的一般条件只是要求它们是正有理数.显然丢番图开始要求解为整数只不过是为了使引导性的问题更容易些.从此以后,"数"这个字应解释为"有理数".

问题 Ⅰ-28 求两个数,使它们之和及它们的平方之和各为给定的数.

求解这个问题我们要加一个必须满足的条件,即平方之和减去和的平方应为一平方数.在上述问题中,假定和为20,平方和为208.

这个问题的一般形式为:$x+y=a, x^2+y^2=b$,这是巴比伦人解过的一种类型的问题.还有三种巴比伦人解过的类型分别出现于 Ⅰ-27, Ⅰ-29 和 Ⅰ-30,即分别为 $x+y=a, x\cdot y=b$; $x+y=a, x^2-y^2=b$ 以及 $x-y=a, xy=b$.丢番图对上述问题(对其他问题也一样)的解法是严格代数的,而不像巴比伦人的解是拟几何的.他设两个未知量为 $10+z$ 或 $10-z$.带有平方的方程就成为 $200+2z^2=208$,从而得 $z=2$,而要求的两个数为12和8.丢番图的方法可应用于这种形式的任意方程组,翻译成现代语言就是

$$x=\frac{a}{2}+\frac{\sqrt{2b-a^2}}{2}, \quad y=\frac{a}{2}-\frac{\sqrt{2b-a^2}}{2}.$$

他所加的必要条件就是为了保证解为有理数所必需的.

有趣的是,问题27、29和30的解也是12和8,这提醒我们:无独有偶,巴比伦人也总喜欢让一系列的相关问题有相同的解.

丢番图是不是从欧几里得的《数据》(Data)学来他的方法?他是否取用过巴比伦人的材料?这些问题无法回答.然而,明白无误的是在丢番图的解法中没有系统的几何方法论.可见在当时,巴比伦人的代数方法,抛开其几何起源不谈,已经为希腊世界所熟知.

问题 Ⅱ-8 将一给定数分成两个平方数之和.

> 设要将16分成两个平方数之和,令第一个平方数 $=x^2$,则另一个数就是 $16-x^2$,从而要求做到的是使 $16-x^2=$ 一个平方数.我们把这个平方数写成 $(ax-4)^2$,a 为任一整数,4为16的根.例如,令边为 $2x-4$,则其平方为 $4x^2+16-16x$,有 $4x^2+16-16x=16-x^2$.在等式两边加上负项,并将两边相同的项去掉.则得 $5x^2=16x$,从而 $x=16/5$.于是其中一个数为256/25,另一个为144/25,它们的和为400/25或216,每一个都是平方数.[13](图5.4)

这是一个不定方程问题的例子.翻译过来就是一个有两个未知量的方程 $x^2+y^2=16$.这个问题也显示了丢番图的最常用的方法之一.从第2卷开始以后的许多问题中,丢番图要求以二次多项式的形式表示出的解必须是一个平方数.为保证解是有理数,他选择平方数的形式为 $(ax\pm b)^2$,其中 a 和 b 这样来选择,使得二次项或者常数项能从方程中被消掉.在本题中,二次多项式为 $16-x^2$,他选 $b=4$ 和其前的符号为负,所以常数项将被消掉,由此得到的解为正数.剩下的求解过程就一目了然了.对这方程 $x^2+y^2=16$,或者更一般地说,对方程 $x^2+y^2=b^2$ 使用这个方法,你想要多少个解就可以产生出多少个解.取 a 为任意值,并令 $y=ax-b$,则 $b^2-x^2=a^2x^2-2abx+b^2$ 或 $2abx=(a^2+1)x^2$,故得 $x=\frac{2ab}{a^2+1}$.

另一个丢番图需要考虑平方数的例子是:

问题 Ⅱ-19 求三个平方数,使得其中的最大数与中间数之差和中间数与最小数之差的比为给定值.

丢番图假定这个给定的比值为3:1.如果最小的平方为 x^2,则他取 $(x+1)^2=x^2+2x+1$ 为中

QVÆSTIO VIII.

PROPOSITVM quadratum diuidere in duos quadratos. Imperatum sit vt 16. diuidatur in duos quadratos. Ponatur primus 1 Q. Oportet igitur 16 − 1 Q. æquales esse quadrato. Fingo quadratum à numeris quotquot libuerit, cum defectu tot vnitatum quod continet latus ipsius 16. esto 2 N. − 4. ipse igitur quadratus erit 4 Q. + 16. − 16 N. hæc æquabuntur vnitatibus 16 − 1 Q. Communis adiiciatur vtrimque defectus, & à similibus auferantur similia, fient 5 Q. æquales 16 N. & fit 1 N. $\frac{16}{5}$. Erit igitur alter quadratorum $\frac{256}{25}$. alter verò $\frac{144}{25}$ & vtriusque summa est $\frac{400}{25}$ seu 16. & vterque quadratus est.

[illegible]

OBSERVATIO DOMINI PETRI DE FERMAT.

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos & generaliter nullam in infinitum vltra quadratum potestatem in duos eiusdem nominis fas est diuidere cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

QVÆSTIO IX.

RVRSVS oporteat quadratum 16 diuidere in duos quadratos. Ponatur rursus primi latus 1 N. alterius verò quotcunque numerorum cum defectu tot vnitatum, quot constat latus diuidendi. Esto itaque 2 N. − 4. erunt quadrati, hic quidem 1 Q. ille verò 4 Q. + 16. − 16 N. Cæterum volo vtrumque simul æquari vnitatibus 16. Igitur 5 Q. + 16. − 16 N. æquatur vnitatibus 16. & fit 1 N. $\frac{16}{5}$ erit

[illegible]

H iii

图 5.4　丢番图的《算术》1670 年版第 61 页的部分. 该页中包含了问题 Ⅱ - 8 和费马的一条批注,费马在这条注中讲到,不可能将一个立方数分成两个立方数之和,或者更一般地讲,不可能将任一 n 次幂($n > 2$)分成两个 n 次幂之和(来源:Smifason 协会图书馆,照片 No. 92 - 337).

间的平方数. 因为这两个平方数之差为 $2x + 1$,最大的平方数必定是 $x^2 + 2x + 1 + 3(2x + 1) = x^2 + 8x + 4$,为使这个数为平方数,丢番图令它等于 $(x + 3)^2$,在这个情况下选 x 的系数使得 x^2 项可以抵消. 于是有 $8x + 4 = 6x + 9$,因而 $x = 2\frac{1}{2}$,所要求的三个平方数为 $6\frac{1}{4}, 12\frac{1}{4}, 30\frac{1}{4}$. 可是人们注意到,在他选定 $(x + 1)^2$ 作为中间项后,则在 $(x + b)^2$ 中作为整数的 b 只能选择 3 才能得到他所需要的解. 自然,随着初始比值选不同的值,就会有更多的可能解,这正如对第二个平方数作不同的选择会得出不同的解是一样的. 无论如何,在本题中和在所有丢番图问题中一样,只要求一组解.

问题 Ⅱ - 11 引进了另一个普遍的方法,双方程法.

问题 Ⅱ - 11　在两个给定的数上加上同一个(待求的)数,使得这二者都成为平方数.

丢番图取 2 和 3 作为这两个给定的数,如果他所要求的数为 x,就需要 $x + 2$ 与 $x + 3$ 均为平方数. 这样一来他就必需解方程 $x + 3 = u^2, x + 2 = v^2$ 来求出 x, u, v,这又是一个不定方程问题. 丢番图对他的方法是这样来描述的:"取上述两个表达式之差,并将它分解为两个因子,那么,或者(a)

取这两个因子之差的一半的平方,并令它等于较小的那个表达式,或者(b) 取这两个因子之和的一半的平方,令它等于较大的那个表达式".[14]

由于这两个表达式之差为 $u^2 - v^2$,可因子分解为$(u + v)(u - v)$,这两个因子之差为 $2v$,之和为 $2u$.丢番图没有明确提到的是,起初的因子分解必须精心选择,才能使求得的解 x 是一个正的有理数.在本题中这两个表达式之差为 1,丢番图把它分解成$4 \times 1/4$.于是有 $u + v = 4$和 $u - v = 1/4$,所以 $2v = 15/4, x + 2 = v^2 = 225/64$,从而 $x = 97/64$.需指出,如果将 1 因子分解为,例如,$2 \times 1/2$,则不会得出正数解,分解成 $3 \times 1/3$ 也不会.因子分解 $1 = a \cdot 1/a$ 应该这样来选择 a,使得有

$$\left[\frac{1}{2}\left(a - \frac{1}{a}\right)\right]^2 > 2.$$

5.2.2 高次方程

卷 A 的开始有一篇新的引言.由于现在问题开始含有三次甚至更高次幂,丢番图叙述了这些幂的乘法规则:

> 那么当我用 x 来乘x^3 时,结果和我作 x^2 的自乘相同,并将结果称之为 x^4.如果 x^4 被 x^3 来除,结果就是 x,即 x^2 的平方根;如果被 x^2 来除,结果为 x^2;如果被 x 来除,即被 x^2 的平方根来除,结果为 x^3….当 x^5 乘以 x 时,结果就与 x^3 自乘是一样的,也和 x^2 乘以 x^4 一样,结果称之为 x^6.如果 x^6 被 x,即 x^2 的根,来除,结果为 x^5;如果被 x^4 来除,结果为 x^2;如果被 x^5 来除,结果为 x,即 x^2 的根.[15]

用 x^n 的符号体系来翻译丢番图的乘法规则可能会引起一点误解,因为这套符号使得这些规则几乎是自明的.然而应当记住,丢番图自己的符号可是用 $\Delta^Y, K^Y, \Delta^Y\Delta, \Delta K^Y$ 和 K^YK 来分别表示从 x^2 到 x^6.所以他的读者会读到这样一些命题:ΔK^Y 乘以ς等于 K^Y 的自乘,等于 Δ^Y 乘以 $\Delta^Y\Delta$,而且全都等于 K^YK.

丢番图继续在他的引言中解释道,和前面一样他的方程最终将化为这样的形式:某一幂次的一项等于另一幂次的一项,即 $ax^n = bx^m (n < m)$.然而,现在 m 可以是任一小于 6 的数.为了求解,我们必须利用两边同除以较小幂的法则,最后约化成一个等于一个数的"类"(species),即 $a = bx^{m-n}$.这后一个方程很容易解.丢番图对读者这样写道:"当你已熟悉我所提出的问题后,你就能够解出许多我还没有提出的问题,因为我会向你说明解决一大堆问题所用的方法,并且对其中的每种类型的问题用一个例子来作解释."

作为丢番图使用高次幂的例子,我们来考虑:

问题 A – 25 求两个数,其中一个是平方数,另一个是立方数,使得它们的平方之和是一个平方数.

这个问题的目标是要求 x, y 和 z,使得有$(x^2)^2 + (y^3)^2 = z^2$.可见这是一个有三个未知量的不定方程问题.丢番图令 x 等于 $2y$(这里 2 是随便设的),进行指数运算后就得知 $16y^4 + y^6$ 应该是一个平方数,他把看成是 ky^2 的平方.因此有 $16y^4 + y^6 = k^2y^2, y^6 = (k^2 - 16)y^4$,也就有 $y^2 = k^2 - 16$.可见 $k^2 - 16$ 必定是一平方数.丢番图选了最容易的数,即 $k^2 = 25$,所以 $y = 3$.由此得出所要求的两个数为 $y^3 = 27$ 和$(2y)^2 = 36$.这个解很容易推广.对任何的正整数 a,令 $x = ay$.接着求 k 和 y 使得有 $k^2 - a^4 = y^2$,或使得有 $k^2 - y^2 = a^4$.然而,丢番图已经在问题 Ⅱ – 10 中表明了,我们总能找到两个平方数,使得它们的差为给定.

问题 B – 7 表明,丢番图是知道$(x + y)^3$ 的展开式的.他曾这样写道:"任何时候我们想构造一

个由两个不同项之和作边所成的立方时 —— 只要项数不会使我们犯错误 —— 我们应取这两个不同项的立方,加上每一项的平方与另一项的乘积的三倍."[17]

问题 B－7 求两个数,使得它们之和与它们的立方之和各等于两个给定的数.

这个问题就是要求解 $x+y=a, x^3+y^3=b$.这是一个有两个未知量,两个方程的方程组,是一个确定性的问题.这是"巴比伦"问题 Ⅰ－28 的一个推广,该问题是要求解 $x+y=a, x^2+y^2=b$.丢番图的解法确实推广了他在那个问题中所用的方法.令 $a=20, b=2240$,和先前的做法一样,他令这两个待求的数为 $10+z$ 或 $10-z$.那么第二个方程就成为 $(10+z)^3+(10-z)^3=2240$,或利用前面已经讲到的展开式,就有 $2000+60z^2=2240$,或 $60z^2=240, z^2=4$,即 $z=2$,自然,丢番图给出了保证解为有理数的条件,即 $\frac{4b-a^3}{3a}$ 应为一平方数(这等价于更为自然的条件,即 $\frac{b-2(a/2)^3}{3a}$ 应为一平方数).有趣的是,这里的答案和 Ⅰ－28 题的答案一样,都是 12 和 8.

在通读《算术》时,读者往往不知道下一步会讲什么.有许多不同的问题.有时有几个相似的问题会放在一起,例如,有一组是涉及减法的,而它前面那一组则是涉及加法的.但是读者有时候会疑惑,为什么有些类似的问题又没有包括进来.例如,卷 A 中的头四个问题要求的是(1) 两个立方数,其和为一平方数,(2) 两个立方数,其差为一平方数,(3) 两个平方数,其和为一立方数,(4) 两个平方数,其差为一立方数.这张单子里还缺两个问题.第一个是,求两个平方数,使它的和为一平方数——但这个问题已在问题 Ⅱ－8 中解决了.第二个是,求两个立方数,使其和为一立方数.这后一个问题无法解出,对这不可解问题的叙述,其记载可以追溯到 10 世纪,有可能丢番图也知道这个不可解性.至少他试过去解这个问题而失败了.但在他的著作中他对此什么也没有讲.一个涉及四次幂的类似问题出现在问题 Ⅴ－29 中:求三个 4 次幂的数,使其和为平方数.虽然丢番图解决了这个问题,但他并没有提及不可能求得两个四次幂数,其和为一平方数这件事.人们又再次设想,他可能试过去解这后一个问题而未获成功.

最后,在讨论问题 D－11 时,他终于讨论了一个不可能有解的问题:将一给定的平方数分成两部分,使得其中一部分与这个平方数之和也是一个平方数,而其中另一部分减这个平方数之后也将得出一个平方数.他接着说:"因为不可能找到一个平方数,使得将它分成两部分之后,它与每一部分之和还都是平方数,现在我们转而来讨论某些可以求解的问题."[18]

问题 D－12 将一已知平方分成两部分,使得从已知平方数中减去这两部分之一时,剩下的数(在这两种情况下)仍是一平方数.

为什么前面讲的那个情况不可能解呢?要想解 $x^2=a+b, x^2+a=c^2, x^2+b=d^2$,这就意味着

$$3x^2=c^2+d^2, \quad 或 \quad 3=\left(\frac{c}{x}\right)^2+\left(\frac{d}{x}\right)^2.$$

事实上,不可能将 3 分成两个有理数的平方,这很容易用模 4 同余的概念来证明这一点.丢番图本人在这里和以后都未给出这个证明,可当他在问题 Ⅵ－14 中提出 15 不可能是两个平方数之和时,他倒是讲清了为什么是这样.问题 D－12 的解则是非常简单的.

5.2.3 试位法

在卷四中,丢番图开始采用一种新的方法,它使人想起埃及人的"试位法".

问题 Ⅳ－8 将一个数分别加到一个立方数和加到它的边上去,使第一个和等于第二个和的立方.

这就是要求解 $x^3+y=(x+y)^3$,这是一个不定方程问题,丢番图一开始设 $x=2y$,于是有

$$8y^3 + y = (3y)^3 = 27y^3, \quad 或 \quad y = 19y^3, \quad 或 \quad 19y^2 = 1,$$

但是，丢番图这样写道："19 不是一个平方数"，因此它需要用一个平方数来替代. 折回原路，他注意到，19 等于 $3^3 - 2^3$，而 3 是来自假定 $x = 2y$ 再让系数加 1 得出来的，因此他需要找到两个相继的数 $z, z+1$，使得它们的立方相差一个平方. 于是 $(z+1)^3 - z^3 = 3z^2 + 3z + 1$ 应为一平方数. 利用他常用的求平方数的技巧，他令此差值等于 $(1-2z)^2$. 因而 $z = 7, z+1 = 8$. 回到问题的开始，他重新令 $x = 7y$. 那么就有 $343y^3 + y = (8y)^3 = 512y^3$，或 $1 = 169y^2$. 由此推知，这个被加上去的数 y 为 1/13. 所要求的立方数的边长为 7/13. 人们可能注意到，在他求平方数的过程中，在式 $1 - mz$ 中对 m 的选择，$m = 2$，是惟一能使 z 为整数的选择.

在问题 Ⅳ－31 中，丢番图再次发现，他原来的假设不管用. 但在这里的问题是一个混合二次方程（这是《算术》中第一次出现这样的方程），而这个混合二次方程不会有一个有理数的解.

问题 Ⅳ－31　将单位分成两部分，使得，如果分别在它们上面加上一给定的数，这两个和的积为一平方数.

丢番图令这两个给定数为 3 和 5，再把 1 分成 $x, 1-x$，于是应有 $(x+3)(6-x) = 18 + 3x - x^2$ 为一平方数. 因为他常用的解平方数的技巧在这里都不管用（不论 18 还是 －1 都不是平方数），他试用 $(2x)^2 = 4x^2$ 来作为所求的平方数. 但这样得出的二次方程，$18 + 3x = 5x^2$，"给不出一个有理数的结果来". 他需要用形如 $(mx)^2$ 这样的平方数来代替 $4x^2$，这会得出有理数的解. 于是，由于 $5 = 2^2 + 1$，他注意到，如果 $(m^2+1) \cdot 18 + (3/2)^2$ 是一个平方数，则二次方程就可解了. 这意味着 $72m^2 + 81$ 是一个平方数，譬如设为 $(8m+9)^2$（在此，他常用的技巧取得了成功）. 于是 $m = 18$，并且回到开始，再令 $18 + 3x - x^2 = 324x^2$. 接着他只是直接列出解的结果为 $x = 78/325 = 6/25$，而所要求的数为 6/25，19/25.

虽然丢番图在问题 Ⅳ－31 中并未给出解二次方程的细节，但他在问题 Ⅳ－39 中却给出了这个细节. 他在该问题中陈述很容易翻译成求解方程 $c + bx = ax^2$ 的公式

$$x = \frac{\frac{b}{2} + \sqrt{ac + \left(\frac{b}{2}\right)^2}}{a}.$$

假如我们先用 a 乘整个方程，并求解 ax，则这个公式与巴比伦人的公式是一样的. 这技巧实际上已在巴比伦人的问题中用过. 丢番图十分熟悉这个公式和它的各种变形，以致他在以后的各种问题中都采用过它，不仅用来解二次方程，还用来解二次不等式.

问题 Ⅴ－10　将 1 分成两部分，使这两部分别加上不同的给定数后，所得均为平方数.

在这个问题中有一张图，这种情况在整个著作中只出现过两次（图 5.5）. 丢番图假定那两个给定的数为 2 和 6. 他在图中这样来表示它们和 1：令 $DA = 2, AB = 1$ 和 $BE = 6$. 点 G 是这样来选择的，使得 $DG(= AG + DA)$ 和 $GE(= BG + BE)$ 都是平方数. 由于 $DE = 9$，问题就归结为，将 9 分为两个平方数，使得其中一个介于 2 与 3 之间. 如这个平方数为 x^2，则另一个平方数就是 $9 - x^2$. 和前面那些问题不一样，丢番图不能简单地令 $9 - x^2$ 等于 $(3 - mx)^2$（m 为一待定数），因为他需要让 x^2 满足上述不等式条件. 因此他令 $9 - x^2$ 等于 $(3 - mx)^2$，其中 m 待定. 于是有



图 5.5　丢番图《算术》一书中的问题 Ⅴ－10.

$$x = \frac{6m}{m^2 + 1}.$$

他不是把这个 x 的式子代入 $2 < x^2 < 3$，然后来解一个四次不等式，而是选两个接近 2 和 3 的平方

数,分别为 $289/144=(17/12)^2$ 和 $361/144=(19/12)^2$,然后将 x 的表达式代人不等式 $17/12<x<19/12$. 于是有

$$\frac{17}{12}<\frac{6m}{m^2+1}<\frac{19}{12},$$

其中左边的不等式相当于 $72m>17m^2+17$. 尽管相应的二次方程没有有理数解,但丢番图使用了解二次方程的公式,并指出,由于 $\sqrt{(72/2)^2-17^2}=\sqrt{1007}$ 介于 31 与 32 间,数 m 应选择为 $m\leqslant 67/17$. 右边的不等式则表明有 $m\geqslant 66/19$. 因此丢番图就在这两个极限之间挑选了最简单的 m 值,即 $3\frac{1}{2}$. 因此

$$9-x^2=\left(3-3\frac{1}{2}x\right)^2 \quad 以及 \quad x=\frac{84}{53}.$$

于是 $x^2=7056/2809$,而所要求的 1 的分段值为 1438/2809 和 1371/2809.

丢番图的著作,是古希腊幸存下来的惟一真正有关代数方面的著作,可谓影响深远. 它不仅在近古史中受到关注,而且也为许多伊斯兰的学者所研究,其中包括卡拉季(al-Karajī),它的许多问题都被拉斐尔·邦贝利(Rafael Bombelli)所采用,并发表在他于 1572 年出版的《代数学》一书中,而巴切特(Bachet)的第一个希腊文印刷本出版于 1621 年,皮埃尔·费马(Pierre Fermat)对它作了仔细的研读,使他在数论中获得了许多普遍结果,而丢番图本人对这些却往往只是点到为止. 然而,也许更为重要的是,作为一本代数学方面的著作,它实际上是一本讲分析问题的论著. 譬如说,在解每个问题时开始都是假设解 x 已经找到了,从这个假设继续推导下去直至 x 的数值可以通过解一个简单的方程来确定. 至于问题的综合方面即证明这个答案满足所要求的条件,丢番图则从未给出过,因为这只需作算术计算就够了. 从这点看,丢番图的著作与欧几里得的纯粹综合的著作相比,正好是与之对立的另一个极端.

5.3 帕普斯与分析

尽管所有希腊主要的数学家都用了分析与综合的方法,但就我们所知,直至帕普斯(Pappus)的著作中才有对已发表的工作的方法论的系统研究. 帕普斯于 4 世纪初生活在亚历山大城,是希腊传统的最后一批数学家之一(补遗 5.3). 他对我们已经讨论过的那些人的大大小小的著作都非常熟悉,甚至某种方式上推广了他们的工作. 他最负盛名的著作是他的《全集》,由一组八本论述数学中各种课题的著作组成,可能是在他死后由一个想保存他的论文的编辑者将它们汇集在一起的. 全集中各卷的水平参差不齐,但大部分材料都是对帕普斯从前人著作中收集到的有关某些数学课题的研究的综述.

卷 3 的序言提供了一些与这部著作有关的间接而有趣的信息. 帕普斯的这篇序言实际是致一位女几何学教师的信,他在其中抱怨道:"有些声称最近从你那里学会了数学的人对某些问题的说明却是错的"[19]. 帕普斯说这话的意思是指这些人企图用一些于事无补的方法来解某些问题,例如,只用圆和直线来解两个比例中项的问题. 帕普斯并没有说明他是如何知道这种作图是不可能的. 然而,从他的叙述中我们得知在亚历山大的妇女也卷进了数学[30].

卷 5 是《全集》中最完美的一卷,是讨论等周图形的,而等周图形是指形状不同、周长相同的图形. 他在导论中提出了一个与正文中的纯粹数学相对比的内容,其中他谈到了蜜蜂的智慧:

补遗 5.3

亚历山大地方的数学家是什么人？

拉斐尔(Rafael)的油画雅典学派把托勒密画成一个有着意大利人长像的王子，而最常见的希帕蒂娅的"肖像"，被认为是一个叫做加士帕罗(Gasparo)的画家所画，也把她画成像个意大利人的样子.这毫不奇怪：艺术家们通常总是要用他们的同代人来作很久以前的人物的模特儿.但是我们真正想知道的是，在从1世纪到5世纪这段时期中的亚历山大地方的数学家有多少是希腊人.肯定这些人都是用希腊文写作，并且是亚历山大的知识分子社团的一部分.对希腊化埃及的近代研究多数都认为希腊人的社团和本地埃及人的社团当时是共存的，相互间影响很小.那么我们是否应该得出这样的结论，认为托勒密、丢番图、帕普斯和希帕蒂娅等人在种族上是希腊人，他们的祖先在过去的某个时候从希腊迁移过来，但实际上很少与埃及人往来？

自然，我们对这个问题无法作出肯定的回答.但是对公元头几个世纪的纸草书的研究表明，在希腊人与埃及人之间有相当数量的通婚，主要是希腊男人娶埃及女人为妻.而且我们知道，希腊人的婚约也越来越像埃及人的婚约了.此外，从亚历山大的建成来看，一小部分埃及人也被接纳为城市的特权阶级，担任许多市政职务.自然，在这种情况下的埃及人要"希腊化"，遵守希腊的习俗，采用希腊的语言.考虑到我们所提到的这些亚历山大数学家活跃于该城建立几百年以后的时期，那么看来至少可以认为：从种族上来讲，他们属于埃及人或是希腊人的可能性是相等的.无论如何，在没有可靠的依据下，把他们画成纯粹欧洲人的样子是毫无道理的.

毫无疑问，[蜜蜂]相信它们自己是受托来完成这样的任务，就是从上帝那里把这种形式的美食带给较为开化的一部分人类……，认为把这种食品漫不经心地倾洒在地上、树木或者其他不相宜和不规则的物质上是不恰当的.但是，在收集生长在土地上最甜蜜的花朵中的精华部分时，它们准备了盛装蜂蜜的容器，叫做蜂房，[每格蜂房]都相似和大小相等，互相邻接，形状为六边形.

我们于是可以认为，它们是根据某种先验的几何思想来进行设计的.它们肯定会认为，这些蜂房的每一格都应该相互邻接，有共同的边，以免有异物落入空隙，从而把它们的劳作弄脏.能够满足这个条件的直边图形只有三种，我是指规则的图形，即等边和等角的图形，因为不规则的图形会惹这些蜜蜂生气……，[这种图形有]三角形、正方形和正六边形，蜜蜂以它们的智慧为它们的产品选择了那个角度最多的六边形，预见到它比其他两种形状能存更多的蜜(图5.6).



图5.6 卢森堡邮票上的六边形蜂房.

可见，蜜蜂知道这个对它们有用的事实，即消耗相同的材料建造蜂房时，这种正六边形的蜂房比正三角形和正方形蜂房能贮存更多的蜜.但是我们，自认比蜜蜂拥有更多的智慧，将要研究一个更一般的问题，即在所有等边，等角和有相同周长的平面图形中，角越多的，它的面积也越大，其中最大的就是具有相同周长的圆.[21]

5.3.1 分析荟萃

然而，帕普斯《全集》中最有影响的一卷是卷7，"论《分析荟萃》"，其中包含了对从希腊时代以来的分析方法(即希腊人解决问题的方法论)的清晰的讨论.其中心思想在卷7的引言中是这样表述的：

……整体来讲,《分析荟萃》是一部特殊的资料……,是为了满足那些想获得几何学知识,并能够解决别人向他们提出的问题,而且也只对此有用.它由三个人撰写,他们是《原本》的作者欧几里得,波哥(Perga)地方的阿波罗尼奥斯,和老亚里斯泰斯(Aristaeus the elder),并由分析和综合方法开始论述.

这样,分析就是这样一条道路,它从人们想寻求的结果开始,假定这个结果是成立的,通过其推论达到已由综合所确定的某种结论.……有两种分析:一种是寻求真理的分析,叫做"理论性的"分析;另一种是寻求所需要的结果的分析,叫做"问题型"分析.在理论性的分析中,我们假定所探求的是真实的事实,则它所导出的结果根据假设也应该是真的事实,通过这些导出的结果推进到得出某个结论,如果这样得出的结论是真理,则我们开始所探求的结论也是真的,反推分析的过程就是它的证明;但是,如果所得出的结果是错的,则我们开始所探求的结论也是错的.在问题型分析中,我们假设命题为某种已知的结论,则它的推论也应该是对的,通过其推论得出某个结论,如果这所得的结论是可能的和可获得的,是数学家们所谓的"已知的"东西,则我们想求的命题也是可能的,这时证明仍然是这个分析的反推过程;但是,如果我们所得出的结论是不可能的,则这个问题也是不可能求解的.[27]

于是,按照帕普斯的想法,用分析解一个问题或证一条定理,开始就是假设所要寻求的为真,然后由它出发推出一系列推论,直至达到一个"给定的"或我们知道为真的结果.就是说,开始假定所求的 p 为真,然后证明 p 蕴含 q_1,q_1 蕴含 q_2,…,q_n 蕴含 q,这里 q 是某个已知为真的结论.求解一个问题或证明一个定理,正式的综合证明就是反转上述分析的过程,从 q 蕴含 q_n 开始.这个反推的方法常常引起争论;毕竟,不是所有定理的逆定理都能成立.然而事实上,来自欧几里得和阿波罗尼奥斯的大部分重要的定理至少都有部分逆定理成立.这样,这个方法的确常常能提供我们所需要的证明或解.或者,在只有部分逆定理的情况下,也至少说明了,在什么条件下一个问题可以有解.

在现在的文献中,理论性分析的例子很少,这是因为,譬如说,欧几里得从不把他发现证明的方法拿出来与大家分享.但是《原本》卷 XIII 的某些手稿中,却包含了对头五个命题的每一个所作的分析(它们显然是在公元初年由他人加进去的).我们来考虑:

命题 XIII-1 如一直线以中末比分割为两段,则其较大一段与其总长一半之和的平方是一半平方的五倍.

令 AB 在点 C 按中末比划分,AC 为其中较大的一段,再令 $AD = \frac{1}{2}AB$(图 5.7).我们来进行分析.先假设结论为真,即有 $CD^2 = 5AD^2$,并由此来作推论.由于 $CD^2 = AC^2 + AD^2 + 2AC \cdot AD$,因此有 $AC^2 + 2AC \cdot AD = 4AD^2$,但是 $AB \cdot AC = 2AC \cdot AD$,再由于 $AB : AC = AC : BC$,故也有 $AC^2 = AB \cdot BC$.于是 $AB \cdot BC + AB \cdot AC = 4AD^2$ 或 $AB^2 = 4AD^2$,或最后得 $AB = 2AD$,这是已知为真的结果.综合则可通过逆转每一步来进行:由于 $AB = 2AD$,我们有 $AB^2 = 4AD^2$,又由于 $AB^2 = AB \cdot AC + AB \cdot BC$,由此推知有 $4AD^2 = 2AD \cdot AC + AC^2$.在这个等式的两边同时加上 AD^2,得结果 $CD^2 = 5AD^2$.

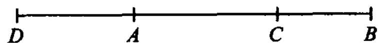


图 5.7 《原本》XIII-1 中的分析.

对希腊数学家来说,问题型分析比理论性的分析更重要.我们已经讨论了这种分析的几个例子,包括三分角的问题及倍立方的问题,以及阿基米德用平面分割球的问题.尽管欧几里得没有这样来表述他的分析,但在解欧几里得 VI-28 的问题时可以用这个方法,在那里的几何代数问题导致要解一个二次方程 $x^2 + c = bx$.在那里的分析表明,为了求解要加上一个附加的条件,即要求有 $c \leqslant (b/2)^2$.

那么,帕普斯一书的卷7可以说是《分析荟萃》的一本指南,后者本身是由好几本几何著作组成,都是在帕普斯前好几个世纪的人写成的.这些著作有,阿波罗尼奥斯的《圆锥曲线》及其它六本书(除一本外均已失传),欧几里得的《数据》及其他两本失传了的书,以及亚里斯泰斯和埃拉托塞尼(Eratosthenes)各写的一本(均已失传),尽管这最后一位作者在帕普斯的引言中并没有被提到.这些著作为希腊数学家们用分析法求解问题提供了必需的工具,例如在处理会引出圆锥截线的问题时,就需要熟悉阿波罗尼奥斯的著作.在讨论可以用"欧几里得的"方法来求解的问题时,《数据》一书中的内容就很关键.

帕普斯的著作本身并不包含《分析荟萃》.它的内容仅仅是为了导读这些著作而设计的.因此,它包括了对这些书的大多数所写的详细的导言,附有大量的引理,其用意是帮助读者读懂原著.帕普斯显然认定这些著作本身对他同时代的大多数读者来说太难了.经过几个世纪教学传统越来越式微了,当时很少有人能像帕普斯那样,还能认识这些百年老书的价值.帕普斯的目标是使更多的人能理解这些经典著作中的数学,办法是帮助他的读者读懂那样一些地方,在这些地方作者只是写一句"显然……!"就一带而过了.他还加进去了一些补充性的结论和一些附加的案例,以及一些不同的证明.

在这些附加的注释中有一个对阿波罗尼奥斯讨论过的三线和四线轨迹问题的推广.帕普斯指出,在原问题中,轨迹是一条圆锥截线.但是,如果超过四条线,则轨迹至今还不甚了了,就是说,"它们的起源和性质还不清楚".对于五线和六线问题的轨迹曲线,至今尚无人能给出作图,他对此感到很失望.五线(或六线)问题是:已知五条(或六条)直线,求一个点的轨迹,使得由与其中三条线相交为已知角的通过它的直线所构成的长方体与由余下的两条直线和某一指定的直线(或由余下三条直线)所作类似直线包含的长方体之比为给定值.帕普斯注意到还可以把这个问题推广到比六条线更多的情况,但在这种情况下,"就不再能说'某个由四条线所包含的图形与剩下的直线所包含的图形之比为给定值'了,因为没有图形有比三维更高的维数".虽然如此,按照帕普斯的说法,我们还可以将这个乘积的比表示成单根线相互间的比值的复合比,这样我们就能讨论直线数为任意多时的问题.但是,帕普斯抱怨说,"[几何学家]决没有[把这个多线轨迹问题]解决到由此确定的曲线可以被识别的程度…….当今研究这些问题的人没有与古人和最优秀的作者同等的素质.看到所有的几何学家都在研究数学的基本原理……同时又为自己也在无为地忙于这些课题而感到羞愧,我已经证明了一些更为重要和有用得多的命题."[23]

帕普斯在卷7的最后叙述了一个他已给出证明的"重要"结果作为该卷的压轴戏,这个结果就是:"旋转体的体积之比就是由旋转的图形之比与这些图形的重心到转轴的类似直线之比的复合比"[24].这个定理的现代说法就是:由一个区域Ω绕一不与Ω相交的轴旋转而形成的立体的体积等于Ω的面积与Ω的重心转过的圆周的周长的乘积.遗憾的是,帕普斯的证明没有记载,有迹象表明,这个证明可能在《全集》的已经失传了的某一卷中.

希腊数学中最清晰的分析方法,大多与我们一般所谓的代数内容有关.选自《原本》的命题XIII-1及VI-28显然是这样的例子.利用圆锥截线的例子在今天我们就会用解析几何去解,这也是大家所熟悉的代数的一个应用.那么,奇怪的是,为什么帕普斯没有提到丢番图的严格的代数著作《算术》并将其作为分析的最佳的例子呢?实际上丢番图著作中的每一个问题都是依照帕普斯的模式来求解的.帕普斯没有将此著作包括进来可能是他认为它与经典的几何著作不在一个水准上.不管怎么讲,并不是纯粹的几何分析,而是丢番图的代数分析和许多其它已提到过的著作中的"拟代数的"分析,才是推动16世纪和17世纪欧洲数学家去进一步阐明代数的概念,并将它发展成为一个可以用来解决甚至是纯粹几何问题的主要工具的动力.[25]

5.3.2 希帕蒂娅和希腊数学的终结

帕普斯想恢复希腊数学活力的目标未能取得成功,部分原因可能是由于不断升级的宗教和政治上的混乱局势影响了亚历山大博物馆和图书馆的稳定.在他那个时代,一度曾受压制的基督教在罗马帝国以一种官方宗教的身份出现.公元 313 年皇帝加勒流斯(Galerius)在东部帝国发布了一道容忍异已的敕令,两年以后康斯坦丁(Constantine)在西部也作了相同的事.实际上康斯坦丁在他 337 年去世以前转而信奉基督教.在六十年的时间里,基督教成了帝国的国教,而古时候对罗马众神的崇拜被禁止了.自然,这种对异教的禁令并不能使每个人都信奉基督教.事实上,在 4 世纪末和 5 世纪初,希帕蒂娅(约 355—415),亚历山大的塞翁的女儿,是该市受人尊敬的优秀教师,不仅教数学,同时还教柏拉图学派的哲学.尽管一直保持着非基督教的信仰,她仍享有学术上的独立性,在她的学生中甚至还有一些著名的基督徒,其中包括昔兰尼(Cyrene)(在利比亚)地方的西奈修斯(Synesius),他后来成了一位主教.

补遗 5.4 希腊数学的衰落

为什么希腊数学在公元前 3 到 4 世纪从它的顶峰这样急剧地衰落下来呢?在对这个问题的各种答案中,最重要的原因就是在于地中海周边地区在宗教领域里的社会政治气候的变化.

对各种古代社会数学发展的考查表明,数学的创造性需要某种智力好奇心的火花来点燃,而仅仅关注实用是不一定能激发这种火花的.但是要想使这种火花传播开来,政府对它的鼓励就不可缺少.巴比伦人应用他们最高级的数学技巧,不是为了日常的生活,而是为了解决有挑战性的智力问题,而政府则鼓励这种研究以帮助训练未来领袖人物的大脑.在希腊文明中,智力的好奇心则更深入人心.在希腊本土,哲学和数学受到广泛的支持.公元前 300 年之后,托勒密家族继续了这种鼓励政策.

但即使在希腊社会,懂理论数学的实际人数也是不多的.总不可能会有很多人能以数学家或天文学家的身份度日,或者能说服统治者给他们提供薪俸.最好的数学家撰写著作,这些著作在各种学校里得到讨论和评注,但并不是什么都能从这些教本中学到.要想维持数学的进步,口头传授也不可缺少,因为在大多数的情况下,光靠自己是很难掌握欧几里得《原本》或阿波罗尼奥斯《圆锥曲线》的内容.这种教学传统如果长期中断,数学研究的整个过程就会受到严重的损害.

有一个因素,如果不是说整个中断了,肯定也是大大削弱了这种教学传统,这就是地方性的政治斗争,它在公元初的前后折腾着东地中海地区.更为重要的原因是,罗马帝国政府显然认为,数学研究不是国家的重大利益,因而不予支持.在罗马对数学学习不加鼓励.很少有希腊学者被邀请来给那些有天份的儿童教数学.不久,在罗马就没有人能看懂欧几里得或阿波罗尼奥斯的著作了,更谈不上进一步发展了.希腊传统在埃及的罗马统治时期也的确继续了几个世纪,这主要是由于亚历山大的博物馆和图书馆还仍然存在,学生们还能继续学习和诠释古代的教本,但是教师越来越少,完成新的著作也越来越少.在 4 世纪末大图书馆的彻底毁坏最终切断了与过去的脆弱联系,尽管雅典和其他地方在一段时间里还有一点有限的数学活动——在这些地方还能找到经典著作的本子——但是把自己的精力贡献给数学的人实在是越来越少,从而传统——包括希腊数学——就寿终正寝了.

尽管有迹象表明在希帕蒂娅之前已有妇女参加到希腊的数学事业中来,但对她们的数学成就

我们所知甚少.关于希帕蒂娅的事迹却有较可靠的资料.她从她父亲那里受到了数学和哲学方面的很全面的教育.尽管保存下来的与她有关的档案只有西奈修斯写给她的一些向她请教科学问题的信件,但是近年来对希腊的、阿拉伯的及中世纪的拉丁文原稿所作的仔细研究可以得出这样的结论,即她是许多数学著作的作者,这些著作包括:她父亲对托勒密《大成》所作的注释的若干部分;阿基米德的《圆的测量》的编辑本,后来的大多数阿拉伯和拉丁文译本都是以此为蓝本的;一本论面积和体积的著作,其中对阿基米德的相关工作进行了重新加工;还有一本与帕普斯《全集》卷5中等周图形有关的教科书.[26] 她还为阿波罗尼奥斯的《圆锥曲线》以及(这在前面已提到过)丢番图的《算术》作过评注工作.

不幸的是,尽管希帕蒂娅在亚历山大有许多有影响的朋友,包括罗马省长奥热斯特斯,但他们主要都是来自上层阶级,而在最高主教西里尔与奥热斯特斯争夺对该市的控制权的斗争中,下层民众基本上都支持西里尔.因此当西里尔散布流言,说这个著名的女哲学家的哲学、数学和天文学研究工作实际上有一部分是巫术时,就有一群人跳出来要消灭这个"撒旦式"的人物,希帕蒂娅的生命就这样遭残害而过早地离开了人世.她的死实际上标志着亚历山大希腊数学传统的终结.

习　题

选自尼可马科斯的《算术入门》

1.设计出一个用来算第 n 个五边形数和第 n 个六边形数的公式.

2.导出底为三角形的棱锥体数的一个代数公式,再导出底为正方形的棱锥体数的公式.

3.证明,在调和比中,两端值之和乘以中值等于两端值乘积的二倍.

4.尼可马科斯定义**次相反比**如下:当三项中的最大项与最小项之比等于两个较小项之差与两个较大项之差的比.证明 3,5,6 成次相反比.再举出了两组成次相反比的三项数.

5.尼可马科斯宣称,如果三数成次相反比,则较大的项与中项之积是较小项与中项之积的二倍;因为他注意 6 乘 5 是 5 乘 3 的二倍.证明尼可马科斯弄错了.

6.尼可马科斯定义了一种"第五种比"如下:当三项中的中项与最小项之比等于这两项之差与最大项与中项之差的比.证 2,4,5 成"第五种比".举出另外两个成这种比的三数组.

选自丢番图的《算术》

7.根据本章开头丢番图的墓志铭确定丢番图的年龄.

8.用 Ⅰ-28 的方法解问题 Ⅰ-27:求两个数使得它们的和与积分别为给定数.丢番图给出的和为 20,积为 96.

9.解丢番图的问题 Ⅱ-10:求两个平方数,使它们的差为一给定数.丢番图令此差为 60.再给出一般的法则用来求给定差为任意值时这个问题的解.

10.推广丢番图对问题 Ⅱ-19 的解,选择任意比值 $n:1$ 和第二个平方数的值 $(x+m)^2$.

11.用双方程法解丢番图的问题 Ⅱ-13:从同一个(待求的)数减去两个给定数,使得减后的结果为两个平方数.(取 6,7 为给定数,然后解 $x-6=u^2, x-7=v^2$.)

12.解丢番图问题 B-8:求两数,使得它们的差和它们的立方的差为两个给定数.(将方程写为 $x-y=a, x^3-y^3=b$.丢番图取 $a=10, b=2120$)确定为了保证解为有理数,a 与 b 所应满足的必要条件.

13.解丢番图问题 B-9:将一给定数分成两部分,使得它们的立方之和是它们之差的平方的给定倍数.这时方程为 $x+y=a, x^3+y^3=b(x-y)^2$.丢番图取 $a=20$ 及 $b=140$,并注意,有解的必要条件为 $a^3\left(b-\frac{3}{4}\right)$ 是一平

方数.

14. 解丢番图问题 D - 12:将一给定的平方数分成两部分,使得一给定的平方数分别减去这两部分时,减后的结果均为平方数.注意,这个问题的解可以直接从问题 Ⅱ - 8 推得.
15. 解丢番图问题 Ⅳ - 9:将同一数分别加到一个立方数和它的边上,使得第二个和为第一个和的立方(方程为 $(x+y)=(x^3+y)^3$. 丢番图在开始时假设 $x=2z, y=27z^3-2z$).
16. 解丢番图问题 Ⅴ - 10,假设两给定数为 3,9.
17. 《算术》的卷六讨论毕达哥拉斯的三数组.例如,解问题 Ⅵ- 16:求一直角三角形,各边为整数,使得一锐角的平分线的长度也是一个整数.提示:应用《原本》Ⅵ - 3 的结论:一三角形一个角的平分线与对边的交点将该边分成两段,这两段之比等于其余两边之比.

与帕普斯《全集》有关的问题

18. 写出《原本》Ⅵ - 28 命题的分析:将等于一给定直线形的平行四边形贴合到一给定的直线上去,但要再减去一个与给定平行四边形相似的平行四边形.只需考虑平行四边形全为直角的情形.首先假定这样一个矩形已经作出来了,再导出以下的条件:"给定的直线形必须不大于以给定直线的一半为边且与减去的图形相似的矩形."
19. 写出《原本》ⅩⅢ - 4 的分析:如一直线按中末比划分,则直线全长的平方与较短段的平方之和等于较长段的平方的 3 倍.
20. 写出五线问题中的轨迹方程.为简单起见,假设所有的直线相互间不是平行,就是正交,这样所有给定角均为直角.
21. 试证一给定周长的六边形,其面积大于相同周长的正方形的面积.
22. 利用帕普斯定理求一锚环的体积.假设这个锚环是由一半径为 r 的圆盘绕一轴旋转而成,这根轴离圆盘中心的距离为 $R>r$.

选自《希腊代表作选集》(约公元 500 年)

23. 解警句 116:母亲,你为什么为了核桃追打我?漂亮的女孩们已经把这些核桃全分光了.因为美利新从我这儿拿走了它们的七分之二,蒂坦拿走了十二分之一,调皮的阿斯提奥西和菲林娜各拿走了六分之一和三分之一.特提丝抢走了二十个,蒂丝比拿走了十二个,瞧,在那里甜蜜地笑着的格劳丝手中拿着十一个.留给我的就剩这一个核桃了.原来有多少核桃?[27]
24. 解警句 130:在四根输水管中,第一根输满一整水池要一天,第二根要两天,第三根要三天,而第四根要四天,如四根同时输要几天?(注意,这个问题与第 1 章习题中选自《九章算术》的一道题相似.)
25. 解警句 145:A:给我十个硬币,那么我拥有的硬币就是你的三倍;B:而如果我从你那里取来相同数目的硬币,我拥有的就是你的五倍.问他们每个人各有多少硬币?

讨论题

26. 为什么罗马时代的测量师不知道三角学,而且在他们的工作中没有用到它?
27. 阅读尼可马科斯命名比率的方案,设计一个词语"公式"来确定任一给定整数比的名称.
28. "对于解决管理罗马帝国的问题来说,二次方程毫无用处".试给出赞同和反对这句话的论据.
29. 在一特定的文明中,影响数学发展的因素是什么?试从已经学过的文明中举例论述之.
30. 为什么在希腊时代问津数学的妇女这么少?

文献和注解

托马斯·希思的 *A History of Greek Mathematics* 和 B.L.范·德·瓦尔登的 *Science Awakening*(在第2章中已引用过)这两本书中均有章节论及本章所讨论的内容.尼可马科斯主要著作的译文可在下书中找到:M.L.D'Ooge,E.E. Robbins,and L.C.Karpinski,*Nicomachus of Gerasa:Introduction to Arithmetic*(New York,Macmillan,1926).这一译文也可从 Great Books 卷 11 中找到,现存的六卷希腊文丢番图著作可在下书中找到:Thomas L.Heath,*Diophantus of Alexandria:A Study in the History of Greek Algebra*(New York Dover,1964).不过希思并不是逐字逐句翻译丢番图的著作,他只不过写出了丢番图论述的大意.对某些问题更接近于从字面上译出者可在下书中找到:Thoas,*Selections Illustrating the Hisfory of Greek Mathematics*.丢番图《算术》中新发现的四卷的译本和评注在 J.Sesiano 所著之 *Book* Ⅳ *to* Ⅷ *of Diophantos' Arithmetica in the Arabic Translation of Qustāibn Lāgū*(New York,Springer,1982)中可找到.J.D.Swift 的论文"Diophantus of Alexandria",*American Mathematical Mowthly* 63(1956),163 - 170,对丢番图的工作有一简短的扫描.帕普斯的《全集》的现存全部内容已译成了法文,见 Paul Ver Eecke,*Pappus d'Alexandria,La Collection Mathematque*(Paris:Desclée,De Brouwer et Cie,1933).对该书卷七的最近英译本,附加了评注,见 Alexander Jones,*Pappus of Alexandria:Book7 of the Collection*(New York,Springer,1986).最近出了一本希帕蒂娅的传记:Maria Dzielska,*Hypatia of Alexandria*,F.Lyra 译(Cambridge:Harvard University Press,1995).尽管该书对她的数学工作讨论得很少,但这个空白已由 Michael A.B.Deakin 的论文"Hypatia and her Mathematics,"*American Mathematical Monthly* 101(1994)234 - 243 所填补.

1. W.R.Paton 译,*The Greek Anthology*(Cambrige:Harvard University Press,1979). 卷五,pp.93 - 94(Book XIV,Epigram126).
2. Edward Gibbon,*The Decline and Fall of the Roman Empire*(Chicago:Encylopedia Britannica,1952)(*Great Books edition*),第四十七章,p.139.
3. Cicero,*Tusculan Disputations*(Cambridge,Harvard University Press,1927)Ⅰ,2.
4. Vitruvius,*On Architecture*(Cambridge,Harvard University Press,1934)Ⅰ.
5. 对罗马测量术一个很好的讨论可见 P.A.W Dilke,*The Roman Land Surveyors:An Introduction to the Agrimensores*(Newton Abbot:David-Charles,1971).
6. Nicomachus,*Introduction to Arithmetic*,Ⅰ.Ⅳ.2.
7. Nicomachus,*Introduction*,Ⅱ.ⅩⅣ,1.
8. Thomas,*Selections Illustrating*,Ⅱ,pp.519 - 523.
9. 这一论述的细节见 J·Sesiano,*Books IV to VII of Diophantos'*,pp.71 - 75.
10. Thomas,*Selections Illustrating*,p.528.
11. Thomas L.Heath,*Diophantus of Alexandria*,pp.130 - 131.
12. 卷 Ⅰ - Ⅵ 的问题摘自 Heath,*Diophantus*;卷 A - D 的问题则摘自 Sesiano,*Books* Ⅳ *to* Ⅶ *of Diophantos*.
13. Thomas,*Selections Illustrating*,Ⅱ p.553.
14. Heath,*Diophantus*,p,146.
15. Sesiano,*Books* Ⅳ *to* Ⅶ *of Diophantos'*,p.88.
16. 同上,p.87.
17. 同上,p.130.
18. 同上,p.165.
19. Thomas,*Selections Illustrating*,Ⅱ.p.567.
20. 认为 Pandrosian 是个妇女的论据见 Jones,*Pappus of Alexandria*.
21. Thomas,*Selections Illustrating*,Ⅱ,pp.589 - 593.
22. 这段译文摘自 Michael Mahoney,"Another Look at Greek Geometrical Analysis",*Archive for History of Exact Sciences*

5(1968),318 – 348,以及 Jones, *Pappus of Alexandria*. Mahoney 关于希腊分析的某些结论引起了争议,见 J · Hintikka 与 U. Remes, *The Method of Analysis: Its Geometrical Origin and Its General Significance*(Boston, Reidel, 1974).

23. Jones, *Pappus of Alexandria*, pp. 120 – 122,以及 Thomas, *Selections Illustrating*, Ⅱ. p. 601.

24. Jones, *Pappus of Alexandria*, p. 122.

25. 在 J. Klein, *Greek Mathematical Thought and the Origin of Algebra*(Cambridge, MIT Press, 1968) 中对丢番图的代数分析及其对代数发展的影响有广泛的讨论.

26. 把各种数学著作归之于希帕蒂娅的详情见 Wilbun Knorr, *Textual Studies in Ancient and Medieval Geometry*(Boston, Birkhauser, 1989) 以及 Michael Deakin,"Hypatia and Her Mathematics."

27. 这个问题和接下去的两个问题均采自 Paton, *The Greek Anthology*.

希腊代数和分析概览

公元 1 世纪末	尼可马科斯(Nicomachus)	初等数论
公元 3 世纪中	丢番图(Diophantus)	不定方程
公元 4 世纪初	帕普斯(Pappus)	希腊数学总结
约公元 355—415 年	希帕蒂娅(Hypatia)	对阿波罗尼奥斯、丢番图著作的评注

第二篇

中世纪的数学

500—1400

第6章 中世纪的中国和印度

> 数学是一门非常重要的科学. 朱世杰的《四元玉鉴》对世界数学做出了巨大贡献. 谁能很好地利用这本书, 谁就能掌握格物致知之学, 治国平天下之道. 难道不应该仔细研究并从中学到些什么吗?
>
> ——《四元玉鉴》的介绍[1]

一本波斯著作中讲了一个数学, 占星家预言印度数学家巴什迦罗(Bhāskara, 1114—1158) 的女儿丽拉沃蒂(Līvāvatī) 将终身不嫁. 作为天文学家和占星家的父亲, 要为他的女儿预卜良辰. 他把一只杯子放在水中, 杯底有小孔, 水从小孔慢慢渗入, 杯子一旦沉没, 便是佳期降临之时, 丽拉沃蒂带着好奇之心观看这只杯子时, 她颈项上一颗珠子落入杯中, 正好堵塞了漏水的小孔, 中止了杯子的下沉, 她"命中注定"终身不嫁. 为了安慰女儿, 巴什迦罗把他的主要著作《天文系统极致》(Sinddhāntaśiromani) 的算术篇用他女儿的名字命名.

6.1 中世纪的中国数学简介

3世纪初汉代灭亡, 中国分裂成几个国家, 这种战乱局面持续到581年. 隋代建立, 37年后进入唐代. 经过大约300年又开始了分裂局面. 宋代(960—1279) 时中国大部分领土得以统一, 随后被成吉思汗领导下的蒙古人推翻. "中国" 数学指的是这3000年中的数学. 尽管其间有无数战乱和朝代的更迭, 但在亚洲东部的这部分地区, 统一语言和统一价值观下的中国文化一直在不断发展.

汉代已建立起文官考试而非家族传承的制度. 皇家考试制度持续至20世纪初. 中间曾经历几次短期的中断. 尽管考试主要内容是中国文学经典, 但为了满足国家管理事业, 诸如观测、税收、编制历法的需要, 还要求官员们具有某种数学才能. 因此朝廷鼓励实用数学的学习, 本章开篇引言中就有体现. 事实上, 各朝代均有皇家数学机构, 训练官员们的"实用数学" 技能. 有的朝代中, 数学是天文机构或记录机构的一部分. 总的说来, 这些机构只研究带有解法的问题集形式的数学著作, 很

少提出新方法.考试通常要求考生复述数学著作的相关内容,使用书中相同的模式解题.并不特别鼓励数学的创造.

然而中国确实出现了创造性数学家,他们不仅将其才能用于改进解决实际问题的老方法,而且还发展出超越实际需要的新方法.本章第一部分就介绍这些中国数学家的工作.我们首先从3世纪数学家刘徽的观测问题开始,然后讨论8世纪从印度传入的解决相关历法问题的重要的三角学方法.接着介绍数学家解决现在称为一次同余式组问题的方法.这些问题最早出现于4世纪的孙子和5世纪的张邱建两人的著作中,但直至13世纪秦九韶才给出完整的解答.最后一部分将讨论13世纪中国数学在方程解法方面的一些著作.除秦九韶的著作外还包括李冶、杨辉和朱世杰的.

6.2 观测的数学和天文学

刘徽(3世纪)生活在汉代灭亡后北方的魏国,关于他的生平我们所知甚少.其主要数学工作是评注《九章算术》.《九章算术》最后几章是基本的测量问题,于是刘徽增加了一个附录处理更加复杂的测量问题.这个附录最后成为一部独立的数学著作,名为《海岛算经》.

这本简短的著作延续了问题集的传统,但只收录了9个问题,每题有问、解、图、注.然而,现存本中只有题目和算法.因为书中没有给出为什么采用这些特殊算法的原因,所以下面的讨论中我们将说明刘徽可能使用的方法.

这部著作以9问中的第1问而得名,问题是怎样算出海岛的高远.另外还有怎样确定树的高度、山谷的深度、河流的宽度.海岛题为:"假设测量海岛,立两根表高均为5步,前后相距1000步.令后表与前表在同一直线上.从前表退行123步,人目着地观测到岛峰.从后表退行127步,人目着地观测到岛峰.问岛高多少,岛与前表相距多远?"[2]

刘徽的答案为岛高1255步,岛与(前)表相距30750步.他还给出了解法(图6.1):

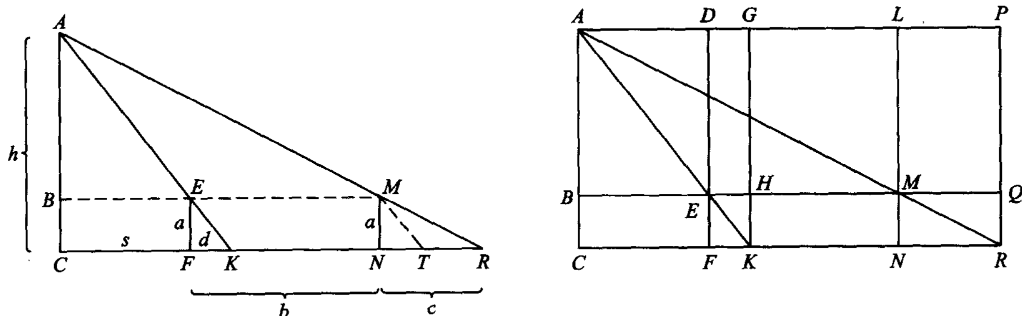


图6.1 《海岛算经》问题1.

用表高去乘两表间距作为被除数(实),用观测点距两表距离之差作除数(法),两数相除,所得之数加表高得岛高.(岛高 h 用公式 $h = a + \dfrac{ab}{c-d}$ 表示,a 为表高,b 为两表间的距离,c 与 d 是观测点分别距前表和后表的距离.)求前表与岛的距离,用前表退行之数乘表间距作为被除数"实",观测点距两表的距离之差为除数"法",两数相除得岛与前表的距离.(距离 $s = \dfrac{bd}{c-d}$.)[3]

因为解题过程中使用了两个差数,所以刘徽将他的方法称为"重差"(double differences),现代数学中用相似三角形:作 MT 平行于 EK,则 $\Delta AEM \sim \Delta MTR, \Delta ABM \sim \Delta MNR$,于是 $ME : TR = AM : MR = AB : MN$.所以

$$AB = \frac{ME \cdot MN}{TR} = \frac{FN \cdot EF}{TR},$$

岛高 $h(= AB + BC)$ 为:

$$h = \frac{FN \cdot EF}{TR} + EF = \frac{ab}{c-d} + a,$$

正是上文中的公式.论证同样给出了岛与前表距离 s 的刘徽的结果.

还有一种方法可得出刘徽的公式.13 世纪中期,杨辉讨论了这一问题,给出了只使用全等三角形和面积关系的合理解释,这种方法是早期中国的数学技巧.因为三角形 APR 和 ACR 全等,三角形 ALM 与三角形 ABM 全等,不规则四边形 $LPRM$ 与 $BMRC$ 面积相等.从中分别减去全等三角形 MQR 与 MNR 可知矩形 $LPQM$ 和矩形 $BMNC$ 的面积相等.同理,矩形 $DGHE$ 与矩形 $BECF$ 相等,从而矩形 $EMNF$(= 矩形 $BMNC$ - 矩形 $BECF$) = 矩形 $LPQM$ - 矩形 $DGHE$.这些矩形面积可以写成

$$\begin{aligned} FN \cdot EF &= PQ \cdot QM - GH \cdot HE = PQ \cdot RN - PQ \cdot FK = PQ(RN - FK) \\ &= AB(RN - FK). \end{aligned}$$

于是 $AB = \dfrac{FN \cdot EF}{RN - FK}$,岛高 $h = AC$ 可由下式得出:

$$h = AC = AB + BC = \frac{FN \cdot EF}{RN - FK} + EF.$$

岛与前表距离 $s = CF$ 可依据矩形 $DGHE$ 和矩形 $BCFE$ 的面积等量关系得出,即有 $CF \cdot BC = DE \cdot EH$,其中 $DE = AB$.

书中第 4 题刘徽通过沿着深谷谷壁的两次观测计算谷深(如图 6.2).x 为所求谷深,单位为尺.现代方法要使用相似三角形,即有

$$\frac{6}{8.5} = \frac{y + 30}{z} \quad \text{和} \quad \frac{6}{9.1} = \frac{y}{z}.$$

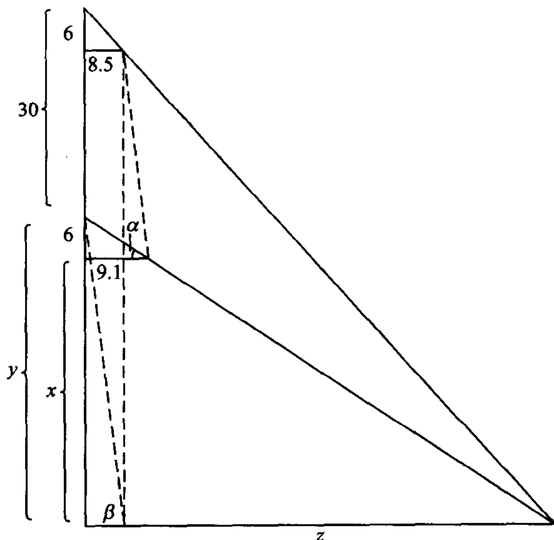


图 6.2 《海岛算经》第 4 题.

可得 $6z = 8.5(y+30) = 9.1y$. 于是 $0.6y = 8.5 \times 30$, $y = 8.5 \times \frac{30}{0.6} = 425$. 刘徽给出了精确的计算,得出谷深 419 尺,比上面所得数字少 6 尺. 他可能使用了与求解问题 1 类似的面积法.

使用相似三角形的计算通常被称作"三角术". 我们可以认为 y 值是由 8.5 乘以图 6.2 中角 β 的正切值 $\left(\frac{30}{0.6}\right)$ 得到.《海岛算经》中别的题目中也同样有长度与角度正切值的乘积. 但因为刘徽及其后来的评注者们都没有提到过这样的角,很难说这部著作中使用了三角术.

8 世纪,中国天文学家在正切表中使用了真正的三角法来计算不同角度的正切值. 同其它地方统治者一样,中国的皇帝对用以预测如日月食等天象的历法问题委感兴趣. 不幸的是中国天文学家由于不完全理解太阳和月亮的运动而不能很准确地预测日月食. 印度天文学家因为受到希腊人创造的几何模型的影响,在这方面较为成功. 于是在 8 世纪,当佛教在印度和中国盛行时,两国有许多僧徒互访,唐朝的皇帝引进了印度的学者和新的技术. 这些学者在瞿坛悉达(Chutan Hsita,8 世纪初)的领导下,在印度典籍的基础上,于 718 年编成一部天文学著作,名为《九执历》(九执指日、月、五星及两颗不可见星). 特别地,这部著作中包含有一张半径为 3438′,间隔 3°45′ 的正弦表.(详见 6.6.1 节.)

724 年,唐朝司天台开展了全国大地测量计划以确定同一子午圈(东经 114°)从 29° 到 52° 纬度范围内冬夏至及春秋分 8 尺高圭表的影长. 著名天文学家一行和尚(683—727)对所测结果进行分析. 一行的目的是使用观测值和各种插值法计算各地影长、昼夜长度和日月食.(一行不知道地球是球状的,因此没有使用古希腊模型.)一行在《大衍历》中构造了晷影表,表中以太阳的天顶距 α 而非以纬度和日期为基础. 一行的表给出了太阳天顶距从 1 至 79 度每隔 1 度的 8 尺圭表的影长值. 按现代术语,这是一张 $s(\alpha) = 8 \cdot \tan\alpha$ 的函数表,是正切表的最早记载.[4]



图 6.3 中国邮票上的一行.

我们不知道一行怎样计算出这张表. 一些学者将这张表与印度天文学著作《九执历》中的正弦表作详细比较后,作出这样的结论,一行使用正弦表作内插法并使用所得结果和公式 $s(\alpha) = 8\frac{\sin\alpha}{\sin(90-\alpha)}$ 计算影长. 尽管《大衍历》和《九执历》都在中国被保留,一行的正切表思想不是中国传统的继承. 三角法在中国也不再出现,直至 17 世纪中国向西方开放. 另外,9 世纪的伊斯兰典籍中出现了晷影(正切)表. 在那个时代,这一思想是否在中亚传播还不清楚.

6.3 不定分析

由于历法问题的需要,中国数学家们考虑解一次不定方程的数学问题. 例如,中国人假定在某一特定时刻"上元",60 天记日周期的首日干支甲子、冬至和新月同时出现. 如果上元之后 N 年时冬至点在 60 干支的 r 日,在新月后的 s 日,那么 N 同时满足同余式:

$$aN \equiv r \pmod{60}, \quad aN \equiv s \pmod{b},$$

其中 a 是回归年值,b 是第一次新月到第二次新月的间隔日数. 然而在现存的古代历法记录中,没有迹象显示中国天文学家怎样解决这样的问题.

6.3.1　中国剩余问题

最早的简单的同余问题出现在数学著作《孙子算经》中.这部著作可能著于公元3世纪后,后来成为文职人员的必修课程.其中包含基本的算术运算方法,也包含下面的例子,我们今天称之为中国剩余问题:

"今有物,不知其数,三、三数之剩二;五、五数之剩三;七、七数之剩二,问物几何?"用现代符号表示,所求数 N 同时满足

$$N = 3x + 2, \quad N = 5y + 3, \quad N = 7z + 2,$$

其中 x、y、z 为整数.上式中 N 相当于满足同余式组

$$N \equiv 2(\bmod 3), \quad N \equiv 3(\bmod 5), \quad N \equiv 2(\bmod 7).$$

孙子给出答案为23,同时给出解法:三三数之余2,写下140.五五数之余3,写下63.七七数之余2,写下30.三数相加得233,从中减去210即得23.孙子进一步解释:三三数之余1,写下70;五五数之余1,写下21;七七数之余1,写下15.若三数相加之和超过106,用105减之即得.[5]

用现代符号表示,孙子明确地指出:

$$70 \equiv 1(\bmod 3) \equiv 0(\bmod 5) \equiv 0(\bmod 7),$$
$$21 \equiv 1(\bmod 5) \equiv 0(\bmod 3) \equiv 0(\bmod 7),$$
$$15 \equiv 1(\bmod 7) \equiv 0(\bmod 3) \equiv 0(\bmod 5).$$

因此 $2 \times 70 + 3 \times 21 + 2 \times 15 = 233$ 满足同余式组.因为105的倍数都能被3、5、7整除,所以233中减去2倍的105即得到最小正整数解.因为孙子只介绍了剩余问题的一种类型,我们不知道他是否发展了寻找模为 μ_i 时余数为1,而模为 $\mu_j, j \neq i$ 时余数为0的整数的一般方法.在这个特殊问题中,这些数通过观察很容易得到,用现代记法,有

$$70 = \frac{3 \times 5 \times 7}{3} \cdot 2, \quad 21 = \frac{3 \times 5 \times 7}{5} \cdot 1, \quad 15 = \frac{3 \times 5 \times 7}{7} \cdot 1.$$

在孙子之后约2个世纪出现的《张邱建算经》(475年)是另一部在唐代科举中使用的数学著作.这部书中包含级数的内容和数值方程的解法,并以首次出现"百鸡"问题而闻名,这一问题也在印度、阿拉伯国家和欧洲的数学著作中以不同形式出现.张的原题如下:"今有鸡翁一,直钱五;鸡母一,直钱三;鸡雏三,直钱一.凡百钱买鸡百只,问鸡翁、母、雏各几何?"[6] 用现代符号表示,x 为鸡翁数,y 为鸡母数,z 为鸡雏数,则三个未知数满足两个方程:

$$5x + 3y + \frac{1}{3}z = 100,$$

$$x + y + z = 100.$$

张给出三组答案:鸡翁4只,鸡母18只,鸡雏78只;鸡翁8只,鸡母11只,鸡雏81只;鸡翁12只,鸡母4只,鸡雏84只.关于解法,他只是说"鸡翁每增4,鸡母每减7,鸡雏每益3".即他指出维持钱数和各类鸡数量关系的变化规律.这个问题可能是用《九章算术》中著名的"高斯消去法"的修正方法解决,从张的描述可得到一般解法

$$x = -100 + 4t, \quad y = 200 - 7t, \quad z = 3t.$$

实际上,张的答案仅是三个值为正数的情形.但我们不能确定张是否使用了这种方法或别的方法.

接下来的几个世纪中有几位作者评注了百鸡问题,但没有人成功地给出方法的合理解释或是一般解法.孙子剩余问题也没有合理解释,尽管8世纪初的一行在历法计算中使用了不定分析,通过解同余式组

$$N \equiv 0(\bmod 1,110,343 \times 60),$$
$$N \equiv 44\,820(\bmod 60 \times 3040),$$
$$N \equiv 49\,107(\bmod 89\,773)$$

来符合几个天文周期.答案为 $N = 96\,961\,740 \times 1\,110\,343$.

6.3.2 秦九韶和大衍术

人物小传

秦九韶(1202—1261)

秦九韶生于成吉思汗征服中国南部时期的四川,他在当时宋朝的首都杭州天文馆(负责历法计算和编制的机构)任职,他写道,“尝从隐君子受数.际时狄患,历岁遥塞,不自意于矢石间,尝险罹忧,荏苒十祀”.为了逃避现实,他用心研究数学,“乃肆意其间,旁诹方能,探索杳渺,……,若其小者,尝设为问答,以拟于用,积多而惜其弃,因取八十一题,厘为九类,立术具草,间以图发之.”[7]《数书九章》中的图是用算筹排列的计算一些题的流程图.

秦曾在不同的地方做官,但他“喜奢好大,嗜进谋身”,为官贪暴.他后来很富有,通过投机取巧,他得到许多土地,建了大量的房子,其中有“漂亮女优居院”.事实上他过了一种骄淫奢侈的生活.[8]

1247年,宋代秦九韶(1202—1261)著《数书九章》,其中提出了解一次同余式组的一般方法.这部著作正如15世纪前的大部分中国数学著作一样,受到古老的《九章算术》的影响.和早期著作一样,秦的《数书九章》是一部问题集.其中许多问题和解法都与古代著作中的相似,却比原来的题目更难一些.其中最重要的成就,是用大衍术解一次同余式组,用现代符号写作

$$N \equiv r_i(\bmod m_i), \quad i = 1,2,\cdots,n.$$

《数书九章》中有10个这种类型的剩余问题,在此我们讨论问题I-4*.如书中所有问题一样,这是一个实用问题,涉及从 A 到 G 七个不同钱库征集日息的问题.每个钱库应交的息款为未知数 N.各库所交息款都应按当地兑换率从过去使用的旧纸币折算,其中 A 钱库可用12钱兑换100钱的旧币,其余各钱库依次减少1钱兑换100钱的旧币,至 G 钱库可以6钱兑换100钱的旧币.问题是以每钱库当地的兑换率清点息款,A 库剩余10钱,D 和 G 余4钱,E 余6钱,其余钱库无剩余.用现代记法,问题为求 N,使之分别满足下列关于 A 至 G 钱库的同余式组

$$N \equiv 10(\bmod 12) \equiv 0(\bmod 11) \equiv 0(\bmod 10) \equiv 4(\bmod 9)$$
$$\equiv 6(\bmod 8) \equiv 0(\bmod 7) \equiv 4(\bmod 6),$$

第一步是约简模,使之两两互素.(《孙子算经》中的例子中模数是完全互素的.)秦氏给出了这一做法的详细程序,即求出模 $m_1, m_2, \cdots, m_n$ 的最小公倍数,再求两两互素的整数 $\mu_1, \mu_2, \cdots, \mu_n$,使其最小公倍数与模的最小公倍数相等,且对每个 i,μ_i 整除 m_i.则当对所有的 i, j,$r_i - r_j$ 可被 m_i 与 m_j 的最大公约数整除时,有 $N \equiv r_i(\bmod \mu_i)$ 等价于 $N \equiv r_i(\bmod m_i)$.(这个条件秦没有说明,但他的所有例子都满足此条件.)在上面例子中,11和7互素,放在一边,$9 = 3^2, 8 = 2^3$.但由于 $12 = 2^2 \times 3, 10 = 2 \times 5, 6 = 2 \times 3$,2和3都出现了高次幂,模10化为5,模12和6都化为1.所得互素的模称为“定数”,即 $\mu_1 = 1, \mu_2 = 11, \mu_3 = 5, \mu_4 = 9, \mu_5 = 8, \mu_6 = 7, \mu_7 = 1$.衍母 θ 等于定数的乘积或原模的最小公倍数,此例中 $\theta = 11 \times 5 \times 9 \times 8 \times 7 = 27\,720$.

* 本书此处对题I-4的释义与《数书九章》的原文有出入,译者略有校正——译注.

第二步，秦用定数除衍母所得称为衍数，我们记作 M_i. 这里 $M_1 = \theta \div 1 = 27\,720, M_2 = \theta \div 11 = 2520, M_3 = 5544, M_4 = 3080, M_5 = 3465, M_6 = 3960, M_7 = 27\,720$，对 $j \neq i$ 的每一 M_i 满足 $M_i \equiv 0(\mathrm{mod}\ \mu_i)$.

第三步，秦从衍数中减去每个定数尽可能多的倍数以求出 M_i 对模 μ_i 的余数. 这些余数记为 P_i，为 $P_1 = 27\,720 - 27\,720 \times 1 = 0, P_2 = 2520 - 229 \times 11 = 1, P_3 = 4, P_4 = 2, P_5 = 1, P_6 = 5, P_7 = 0$. 当然，对每个 $i, P_i \equiv M_i(\mathrm{mod}\ \mu_i)$，$P_i$ 和 μ_i 互素且 $P_i \neq 0$.

最后是求解同余式组，特别是求解 $P_i x_i \equiv 1(\mathrm{mod}\ \mu_i)$. 一旦完成这一步骤，变为互素模的同余问题，就类似于孙子的问题，容易由 $N = \sum_{i=1}^{n} r_i M_i x_i$ 求得一个答案. 因为每个 μ_i 整除 θ，从 N 中减去 θ 的倍数得到其余的答案.[9]

为了求解 $P_i x_i \equiv 1(\mathrm{mod}\ \mu_i)$，其中 P_i 和 μ_i 互素，秦所用的方法称为"大衍求一术". 实际上，这个程序相当于欧几里得算法. 秦用算板来图解描述. 这个特殊的同余式组可用经验方法求解，但这里我们用秦文中 $65x \equiv 1(\mathrm{mod}\ 83)$ 为例来说明求一术. 秦在四方形算板的右上角放置 65，右下角放置 83，左上角放 1，左下角不放数字. 他写道："首先用右上行除右下行，所得商乘左上行加入左下行，（同时右下行用相除后的余数替换.）然后用右边上行和下行以少除多，递推互除，所得商数随即递互累乘加入左边上下行…… 直至右上行变为1时停止. 于是左上行的结果即为所求解".[10] 图6.4的图解代表了下列计算：

$$
\begin{aligned}
83 &= 1 \cdot 65 + 18, & 1 \cdot 1 + 0 &= 1, \\
65 &= 3 \cdot 18 + 11, & 3 \cdot 1 + 1 &= 4, \\
18 &= 1 \cdot 11 + 7, & 1 \cdot 4 + 1 &= 5, \\
11 &= 1 \cdot 7 + 4, & 1 \cdot 5 + 4 &= 9, \\
7 &= 1 \cdot 4 + 3, & 1 \cdot 9 + 5 &= 14, \\
4 &= 1 \cdot 3 + 1, & 1 \cdot 14 + 9 &= 23.
\end{aligned}
$$

1　65 0　83	1　65 1　18	4　11 1　18	4　11 5　7	9　4 5　7	9　4 14　3	23　1 14　3

图 6.4　秦九韶算法解 $65x \equiv 1(\mathrm{mod}\ 83)$ 的算板图解.

右行的每个数字是累次替换得到的 65 的系数的绝对值. 即 $18 = 83 - 1 \cdot 65$，替换为 $11 = 65 - 3 \cdot 18$ 得 $11 = 65 - 3 \cdot (83 - 1 \cdot 65) = 4 \cdot 65 - 3 \cdot 83$，而 4 是右行的第二次计算结果. 同样，$7 = 18 - 1 \cdot 11 = (83 - 1 \cdot 65) - 1 \cdot (4 \cdot 65 - 3 \cdot 83) = 4 \cdot 83 - 5 \cdot 65$. 最后结果是 $1 = 23 \cdot 65 - 18 \cdot 83$，因此 23 为同余式的解.（秦总是调整问题以使最后系数是正的.）

在前面的例子中，$P_i \neq 0$ 时同余式组的解 x_i 为 $x_2 = 1, x_3 = 4, x_4 = 5, x_5 = 1, x_6 = 3$. 这些解称为"乘率". 取 $x_1 = x_7 = 0$，秦将每个 x_i 及相应 M_i、r_i 相乘得总数 $r_i M_i x_i$ 完成了求解. 在此例中，仅当 $i = 4$ 和 $i = 5$ 时有非零值：$r_4 M_4 x_4 = 4 \times 3080 \times 5 = 61\,600, r_5 M_5 x_5 = 6 \times 3465 \times 1 = 20\,790$.

其和 82 390 中减去 2 倍的“衍母”或 55 440，得最后结果 $N = 26\,950$.

我们给出的秦的程序概述来自于他著作中的详细描述.[11] 在《数书九章》实际解决的 10 个剩余问题中，秦总是用各种方法修正步骤，不仅为了简化问题，有时也为了证实他能处理大数. 不奇怪，有两个问题是关于寻求几个不同天文周期事件的公共周期的历法问题. 尽管这些问题的数字非常大，还包含分数和小数，秦也能修改他的解法以找出合适的解.

6.4 解方程

《数书九章》中除了首次描述了求解一次同余式组的程序外，还包含许多解多项式方程的问题. 解方程的方法在其它的中国古典著作中都出现过，在《九章算术》第 1 章中详细讨论了求解完全平方方程（$x^2 = a$）的一般解法. 其中还举出混合二次方程的例子，但没有说明解法.

在随后几个世纪中的其它著作中也出现了类似的问题. 例如，在《张邱建算经》中有下面的问题：假设分割圆得弦长 $68\frac{3}{5}$，面积为 $514\frac{32}{45}$，求弓形的高. 答案为 $12\frac{2}{3}$，但手稿中未出现解法. 我们推测作者使用了公式 $A = \frac{1}{2}h(h + c)$ 并将其转化为关于 h 的二次方程. 此例中，去分母后可得方程 $45h^2 + 3087h = 46\,324$. 三次方程出现在王孝通（7 世纪早期）的著作中，但除了有可能采用立方根求根法外仍然没有给出解法. 很明显，在公元一千年中存在一种求解这类方程的方法.

在 11 世纪中期贾宪所著的一本已失传的著作中，作者不仅利用今天我们称作“帕斯卡三角”（Pascal triangle）的数字排列图将《九章算术》中开平方、开立方的程序推广到了开高次方的情形，而且还将这种方法发展、改进为可以求解任意次多项式方程的方法. 杨辉 1261 年的著作中讨论了贾宪的方法.

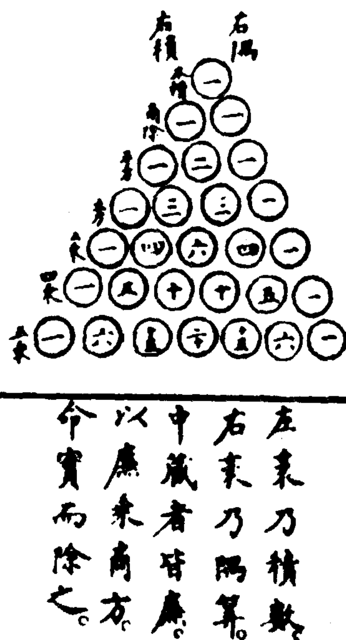


图 6.5　杨辉给出的帕斯卡三角. [资料来源：兰丽蓉. “The Chinese Connection between the Pascal Triangle and the Solution of Numerical Equations of Any Degree”, *Historia Mathematica*, Vol. 7, No. 4, November 1980.]

贾宪的求解二次和三次根的算法基本思想分别源于二项展开式 $(r + s)^2 = r^2 + 2rs + s^2$ 和展开式 $(r + s)^3 = r^3 + 3r^2s + 3rs^2 + s^3$. 例如考虑方程 $x^3 = 12\,812\,904$ 的解，它的根是一个以 2 开头的三位数，也就是说，最有可能的整数解为 $x = 200 + 10b + c$，暂不考虑 c，我们要找出最大的 b 以使

$$(200 + 10b)^3 = 200^3 + 3 \cdot 200^2 \cdot 10b + 3 \cdot 200 \cdot (10b)^2 + (10b)^3 \leqslant 12\,812\,904.$$

依次用 $b = 1, 2, 3, \cdots$ 去试，可知 $b = 3$ 是满足不等式的最大值. 因为

$$3 \cdot (1\,200\,000 + 60\,000 \cdot 3 + 1000 \cdot 3^2) = 4\,167\,000,$$

从 4 812 904 中减去 4 167 000 后又得到一个关于 c 的不等式：$c \cdot (3 \cdot 230^2 + 3 \cdot 230c + c^2) \leqslant 645\,904$. 此例中可知 $c = 4$ 是满足等式的解，于是原方程的解为 $x = 234$.

贾宪已认识到这种解法程序可根据 $(r + s)^n$ 的二项展开式推广到 n 次根（$n > 3$）的情形. 实际上，据杨辉所说，他不仅写出了 6 行的二项式系数的帕斯卡三角，而且他还给出了此三

角形的构造方法:“两肩数值相加即得下层数值.”[12] 贾宪进一步解释了如何利用二项式系数类似地求出更高阶的根.

显然,贾宪实际上做得更多.他认识到他的方法可用来求解任意次多项式方程,尤其是因为这可以看作是开方法程序的一部分;他同时还认识到在算板上通过二项式系数逐步生成各项乘数比直接使用贾宪三角更简单.

6.4.1 秦九韶和多项式方程求解

秦九韶在《数书九章》中首次详细说明并很可能发展了贾宪解方程的方法.我们以书中一个特殊的方程:$-x^4+763\,200x^2-40\,642\,560\,000=0$ 为例说明.这个方程来自于“尖田求积”这一几何问题(参见练习 6).解方程的第一步与开方相同:首先,确定答案的位数;然后,猜出第一位合适的数值,通过经验或分析,此题的答案应是一个以 8 开头的三位数.秦的做法实际上与三次方程求根算法一样,令 $x=800+y$,将其代入方程得到一个新的关于 y 的方程,该方程的根为一个 2 位数.猜出 y 的第一位数然后重复程序.考虑到中国数码的十进制性质,中国人可以根据需要重复这一算法获得任意精确程度的近似根.确实秦九韶对有些题目给出了精确到一位或两位小数的答数,但在另一些场合,有些非整数答案的余数却用分数来表示.

800
-40642560000
0
763200
0
-1

图 6.6 解 $-x^4+763\,200x^2-40\,642\,560\,000=0$ 的第一步算板构造.

当然,中国人没有使用与 1819 年霍纳所用的类似于现代数学的技巧将 $x=800+y$“代入”原方程.问题列于算板上,每行代表未知量特定的幂次(图 6.6).由于空间的原因,我们水平地写出系数,例如刚才的题目,开始的图式是

$$-1 \quad 0 \quad 763200 \quad 0 \quad -40642560000.$$

设定首次近似根为 800,秦九韶对原多项式用 $x-800(=y)$ 做重复(综合)减根变换.第一步得:

800	−1	0	763200	0	−40642560000
		−800	−640000	98560000	78848000000
	−1	−800	123200	98560000	38205440000

秦九韶算板程序的描述正是说明通过乘上和加上(或减去)什么数来得到第三行的数值.例如,-1 乘 800 的结果与 0 相加,所得(-800)乘 800 再与 763 200 相减.在代数学意义上,第一步表明原多项式被下式所替代:

$$(x-800)(-x^3-800x^2+123200x+98560000)+38205440000$$
$$=y(-x^3-800x^2+123200x-98560000)+38205440000.$$

秦九韶将此程序重复了三次,每次用 $y=x-800$ 除商多项式,最后得:

$$0=-x^4+763200x^2-40642560000$$
$$=y\{y[y(-y-3200)-3076800]-826880000\}+38205440000,$$

或

$$-y^4-3200y^3-3076800y^2-826880000y+38205440000=0.$$

当然,秦九韶的算板上仅有数值.他的每一步的图示在这里被联成一个大图:

800	−1	0	763200	0	−40642560000
		−800	−640000	98560000	78848000000
	−1	−800	123200	98560000	38205440000
		−800	−1280000	−925440000	
800	−1	−1600	−1156800	− 826880000	
		−800	1920000		
800	−1	−2400	− 3076800		
		−800			
800	−1	− 3200			
	− 1				

40	−1	−3200	−3076800	−82688000	38205440000
		−40	−129600	−128256000	−38205440000
	−1	−3240	−3206400	−955136000	0

倒数第三行包含了秦氏关于 y 的方程的系数和他所猜测的两位数根的第一个数字 4(这个值可以简单地用 3820544000 除以 826880000 得到). 在此例中,就像今天著作中一般的例子,答案是"整数". y 的方程恰好被 $y-40$ 整除. 原方程的解为 $x = 840$.

我们以帕斯卡三角形作参考,分析秦氏对贾宪方法的描述以及秦又是如何逐步得到二项式系数的. 下面是用秦氏方法解方程 $x^3 = 12\ 812\ 904$,其解的过程布局如下:

200	1	0	0	−12812904
		200	40000	8000000
200	1	200	40000	− 4812904
		200	80000	
200	1	400	120000	
		200		
200	1	600		
	1			
30	1	600	120000	−4812904
		30	18900	4167000
30	1	630	138900	− 645904
		30	19800	
30	1	660	158700	
		30		
30	1	690		
	1			
4	1	690	138700	−645904
		4	2776	645904
	1	694	161476	0

在这张表中容易看出二项式系数. 例如,第 9 行表示第 2 位 b 的方程 $(10b)^3 + 3 \cdot 200 \cdot (10b)^2 +$

$3 \cdot 200^2 \cdot 10b + (200^3 - 12812904) = 0$,正如贾宪的说明.

秦氏本人既没有给出他的程序理论的说明,也没有提到巴斯加三角形.但是因为他在《数书九章》中用这种方法解了26个不同方程,而且他同时代的几位数学家用同样的方法解决了相似的方程,很明显,当时中国的数学家已经掌握了解决这类问题的正确算法.因为这种算法在欧洲的出现比秦氏时代晚了5个世纪,所以值得作更多的评论.

首先,书中仅简要说明了如何猜测出方程的根的各位数值.在一些例子中,很明显解题者简要地用常数项与未知数的一次项系数试除,就像通常求平方根算法中所作的一样.有时需要几次试除,而作者选择其一.但一般地,能推测出中国数学家拥有一张广泛的乘幂表可用来做各种猜测.第二,在书中没有提到多根.实际上,上面提到的秦氏的四次方程还有另一个正根240和两个负根.用同样的方法,但第一位试根为2时易得到根240.但在此例中,构成方程的几何问题仅有一个解840,而秦氏并不处理抽象的方程.第三,负数的运算和正数一样容易.中国人使用不同颜色的算筹表示正负两类数,而且很早就提出了正负数的正确算法.另一方面,负根却未出现过,这是因为构成方程的问题都只有正根.第四,因为中国人能处理负数,他们一般用相当于等式 $f(x) = 0$ 的形式来表示方程.这与巴比伦和中世纪伊斯兰国家是完全不同的.最终出现的中国人解二次方程的方法和巴比伦不同,巴比伦人发展了仅能用于解这种方程的公式法,而中国人发展了最终能推广到任意次方程的数值算法.

6.4.2 李冶、杨辉和朱世杰的工作

与秦九韶同时代的三位数学家李冶(1192—1279)、杨辉(13世纪下半叶)和朱世杰(13世纪后期)也在解方程方面做了杰出的贡献.但因为蒙古人和金及南宋之间持续了近一个世纪的战争,这些数学家彼此之间有多大的影响是令人怀疑的.

人物小传

李冶(1192—1279)

李冶生于黄河以北的河北省正定的官宦之家.1230年他通过了文职官员的考试,在北方金王朝中任职.但在几年中,他的管区和整个金国被蒙古军队攻破,因此李冶弃官,用自己的余生做学问.1260年忽必烈登上皇位后,请李冶在蒙古朝廷任职,李冶出任了很短的时间后,于1266年辞官,回到封龙山区隐居.

李冶写了两部重要数学著作,1248年所作《测圆海镜》和1259年写成的《益古演段》,李冶在其他领域也有很多著作.《测圆海镜》讲述直角三角形内接圆的性质,但主要是关于建立和解决代数方程来处理这些性质.《益古演段》相似地处理关于正方形、圆、矩形、梯形的几何问题,但它的主要目的是教人们建立合适的二次方程的方法以解决问题.

我们给出李冶《益古演段》中的一个例子:

问题8:今有方田一段,内有圆池水占之,外计地三千三百步,方圆周长和三百步,问方圆周长各几何.[13]

李冶的讨论与现代的方法在实质上完全相同,他设 x 是圆的直径,$3x(\pi = 3)$ 为周长,那么 $300 - 3x$ 为正方形周长,将它平方后得 $90\,000 - 1800x + 9x^2$ 为16个正方形面积.又因为$\frac{3x^2}{4}$为圆形池塘的面积,$12x^2$ 为16个原池的面积.两表达式之差即 $90\,000 - 1800x - 3x^2$ 等于16份圆池外面积,亦即 $16 \times 3300 = 52\,800$,所需方程为 $37\,200 - 1800x - 3x^2 = 0$.与秦九韶的工作相比,李冶仅断言

20 是根或直径，因此圆周长为 60，正方形周长为 240.

有趣的是，李冶几乎总是在代数推导之后又给出几何推导(图 6.7). 这里大正方形的边长为 300，为所给周长之和. 阴影部分面积为 16×3300. 因为 $300x$ 是每个长条形的面积，x^2 是每个小正方形面积，$12x^2$ 是 16 个圆池的总面积，他构造了方程 $300^2-16\times3300=6\times300x-9x^2+12x^2=1800x+3x^2$，或如前所述的方程 $37\,200=1800x+3x^2$(注意图中右下角代表三个小正方形).

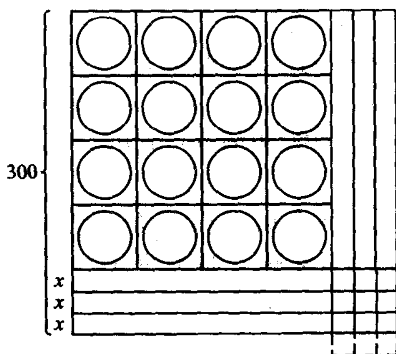


图 6.7 李冶《益古演段》的问题 8.

因此，这部著作提供了中国数学发展的更多证据. 不仅是问题的解法，而且问题最初的提出，都有着几何的基础. 因为数值运算结果都在算板上记录、运算，中国学者们最终认识到算板上的形式并将之转变为数值算法. 同时，他们可能开始将几何概念如正方形等抽象为算板上的一个简单位置，并进一步抽象为未知量平方的概念. 一旦用抽象的符号表示并求解高次方程，那就十分方便. 秦九韶的方程是建立在实际问题，特别是几何问题的基础之上，在解这些问题时他用了不具有几何意义的未知量的幂.

前面已经谈到杨辉的著作中提到了贾宪. 关于杨辉的生平，我们只知道他生活于南宋时期，其它了解甚少. 他的现存的两部著作是《详解九章算法》(1261) 和《杨辉算法》(1275).《杨辉算法》与李冶的著作相似，也包含二次方程的内容，但杨辉较李冶更详细地给出了解这些方程的方法. 总的来说，杨辉使用了与秦九韶解方程相同的方法，但他同时也提出了一种替代方法，这种方法更像我们前面已经讲过的中国人的开平方根的方法，即在推导第二个方程时利用第一次近似的加倍. 另外，杨辉也用含有正方形和矩形的几何图示来说明所用的各种数值方法.

13 世纪中国另一位重要的数学家是朱世杰，我们对他的生平知道很少，他可能生于现在北京的郊区，他一生大多时间从事于数学教育，他于 1299 和 1303 年分别写成《算学启蒙》和《四元玉鉴》.

朱世杰的主要数学贡献之一是把秦九韶解多项式方程的方法发展为解方程组的程序. 这里以《四元玉鉴》中的第 1 题为例说明他的方法. "今有黄方乘直积，得二十四步，只云股弦和九步，问：勾几何?"[14] 即设直角三角形内切圆直径长与两直角边长的乘积为 24，垂直边与弦长之和为 9，求水平边长. 值得注意的是，在中国的表达方式中，由于使用单个汉字代替"直角三角形内切圆的直径长度"及"两直角边乘积"，所以这种表达法相当于我们的符号方程，这里给出的现代翻译保留了中国著作中的特点. 设 a 为垂直边长，b 为水平边长，c 为弦长，d 为圆直径长(图 6.8)，那么此题可翻译为 2 个方程

$$dab=24,$$
$$a+c=9.$$

假定朱世杰还知道下面两个方程

$$a^2+b^2=c^2,$$
$$d=b-(c-a),$$

其中第二个方程给出了直角三角形内切圆直径与三角形边的关系(注意到圆心位于三角形三个角分线上).

遗憾的是也像许多的中国著作一样，这里只有解法的简要说明和关于 b 的 5 次方程，所以我们试图补全被遗掉的部分细节，可能的解题程序如下：首先由 $b^2=c^2-a^2=(c-a)(c+a)$ 和

$(c+a)=9$ 可得，$b^2=9(c-a)$，然后用 9 乘方程 $(c+a)-(c-a)=2a$ 得 $9(c+a)-9(c-a)=18a$，因此有，$81-b^2=18a$ 及

$$18ab=81b-b^3. \tag{6.1}$$

接着用 9 乘 $d=b-(c-a)$ 得 $9d=9b-9(c-a)$ 或

$$9d=9b-b^2. \tag{6.2}$$

方程(6.1)与(6.2)相乘，得

$$162dab=729b^2-81b^3-9b^4+b^5.$$

(这是朱世杰写出的第一个方程.) 因为 $dab=24$，那么他需要解 b 的一个 5 次方程：

$$b^5-9b^4-81b^3+729b^2-3888=0.$$

他只给出结果 $b=3$ 而未给出解法的详细说明.

从书名推测，朱世杰的《算学启蒙》讲述的是基础问题，可能是为初学数学者编写或是算学馆的参考书. 总的说来，它的内容和方法基本上是《九章算术》的重复或适当的修改. 前几章是算术问题，接着是比例、利息、税收、面积和体积问题，其中沿用了许多汉代得出的公式，另外还有用盈不足术及高斯消元法解线性方程组. 相比之下，《杨辉算法》中也有很多解方程的例子以及早期孙子的不定问题、"百鸡问题" 等. 但从其中的内容来看，杨辉在解题方法上并没有比前人有很大的提高，(显然杨辉不知道秦九韶的著作)，杨辉还总结了一些垛积求和问题，而朱世杰对此也有过讨论.

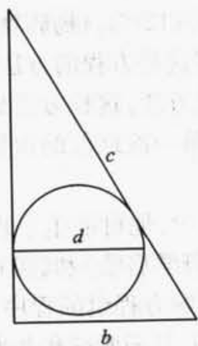


图 6.8　朱世杰《四元玉鉴》第 1 题.



图 6.9　台湾邮票上的利玛窦.

对中世纪中国的数学，我们有如下总的看法：中国数学家精于求解多种代数问题，许多方法起源于对几何问题的思考，后来却变为纯粹的代数程序. 从我们现有的资料来看，也表明中国数学研究最初是源于官府，在后来的几个世纪，计算技巧方面取得了很大的发展，但又由于中国尚古的思想对数学发展的影响，如《九章算术》中的一些错误也被重复了几个世纪，这使得中国数学没有质的飞变，直到 13 世纪，数学家们充分利用算板以及方程中的数字记法，但算板的使用对数学的发展也有局限，因此中国数学未能发展出像西方几个世纪以后发展起来的那种方程理论. 另一方面，当时政治形势也导致蒙古统治时期(元朝) 数学活动的衰落，而 13 世纪一些著名的数学著作都不再被学习，最后在 16 世纪后期，传教士利玛窦(1552—1610) 来到中国后把西方数学传入中国，中国传统数学开始消失.

6.5　中世纪印度数学介绍

公元前 1000 年中叶，一些雅利安部落在印度北部发展壮大，到 6 世纪时，其中较强大的摩揭陀

国占了统治地位,在此或更早些,他们已经建立了等级制度,波罗门教成为印度主要的宗教,它的教义由波罗门教士一代代地口传下来,后来写成文字的形式.为了有助于记忆,它们多被写成诗歌形式,这使得其它非宗教著作也按这种风格写作,例如天文学和数学的著作基本都按这种风格撰写的.印度的天文学与其它国家一样,也是由官方控制.由于缺少书面文献,所以我们对印度数学的发展并不十分清楚.另一方面,对现有的文献,由于它们多为诗词的形式,也不容易把握其内容的准确性.

公元前327年,亚历山大大帝跨过印度库什山进入印度北部,并在随后两年中陆续征服了印度其它小国家.此后,希腊文化开始传到印度.在亚历山大的征程中,随行的有科学家和历史学家,把他们的科学和文化传播到所到之处.亚历山大的一个任务就是"开化"东方.在印度人看来,自己已经是很"文明"了.这样,每个民族都把其它民族看成是"蛮荒".公元前323年,亚历山大去世,他的这一宏图也就随之结束.印度的领土很快被摩揭陀王旃陀罗笈多·毛里亚(Chandragupta Maurya)收复,旃陀罗笈多同亚历山大在西亚的继承者塞琉古建立了友好关系,通过这种关系增进了思想交流.旃陀罗笈多死后不久,阿育王称帝,他继续征服了印度其它的地区,然后皈依佛教并向东方和西方邻国派遣传教者.在全国各地都有刻着阿育王法令的柱子,这些柱子上有印度数码的最早的书面证据.



图6.10 印度邮票,右边是阿育王石柱.

公元1世纪,贵霜人占领北印度,贵霜帝国不久成为罗马帝国与东方贸易的中心.4世纪早期,北印度又在笈多时期重新统一长达一个半世纪,在此期间印度的艺术、医学、开办大学等文化活动达到顶盛.正是在这一时期,印度殖民者将印度文化传播到了东南亚地区,包括缅甸、马来西亚和印度支那.

606年,戒日复兴了北印度王国,但647年他死后,国家逐渐衰落,北印度又分裂为几个小国.这一时期印度南部也是由几个小国组成.虽然在这样的分裂状况下,但印度次大陆基于统一书写语言——梵文——的文化统一性仍然存在.因此,即使是7世纪以后,我们仍可以谈论印度数学.8世纪初,阿拉伯人入侵北印度,伊斯兰和印度间发生了多次大规模的战争.最后在12世纪晚期,穆罕默德·古里(Mohammed Ghori)率领的穆斯林军队统治了北印度,并于1206年建立起德里苏丹国(Moslem Sultanate of Delhi),这个苏丹国延续了300多年,它还曾征服了印度南部的一些国家,而这些国家长期以来一直保持着相对于北方的独立.

在这几个世纪里,虽然有不断的入侵和新国家的建立,但天文研究总是受到鼓励的.无论哪个统治者,都需要天文学家来解决历法问题以及一些占星活动,因此这期间的印度数学著作大多数是天文学著作.与其它国家一样,印度数学也是在研究实用数学的基础上给出了创造性的发展,提出了新的数学.在本章下面的内容中,我们先讨论印度三角学的发展并举出5—7世纪印度著作中的例子来说明有关的方法及技巧,然后是7世纪的波罗摩笈多和12世纪的巴什迦罗Ⅱ的著作中关于不定方程(包括佩尔方程)的解法,最后是中世纪各位印度数学家的代数学和组合学工作.

6.6 印度三角学

公元100年,贵霜王朝和笈多王朝时期,有很多证据可以说明希腊天文学知识可能由罗马贸易通道传入了印度,但奇怪的是托勒密的天文学和数学却没有传入,而托勒密的继承者希帕科斯的著作却传入了印度.由于希腊天文学的需要,三角学得以发展,又由于印度天文学的需要,三角学得以

改进.

6.6.1 正弦表的构造

我们所知印度最早的三角学著作是成书于公元前5世纪的《毗坛摩诃悉昙多》(*Paitāmahasiddhānta*),它也是世界上最早的天文学著作,其中与数学相关的内容流传了几个世纪.为了利用球面三角学知识解决天文问题,《毗坛摩诃悉昙多》给出了一张"半弦表",梵文译为*jyā-ardha*(见补遗6.1).与托勒密求解三角形时使用了倍角所对的半弦值的弦表一样,印度的数学家也觉得用倍角所对半弦值构造数表要比用整弦简单,因而在这部著作和后来的天文著作中都使用这种半弦"函数".托勒密构造的弦表以60为半径,而几个世纪之前的希帕科斯以3438为圆半径.由于后者的半径在《毗坛摩诃悉昙多》中被作为基本半径使用,所以我们推测希帕科斯的三角学较托勒密三角学更早传入印度.为了方便,我们在这里使用现代术语"正弦"而不用"半弦",但要知道直至18世纪,正弦一词才被用来表示特定半径的圆中线段的长度.

我们不考虑保存不完整的《毗坛摩诃悉昙多》中的正弦表构造,而来讨论稍晚的阿耶波多(Āryabhata,生于476年)的《阿耶波多历数书》(Āryabhatīya)中的正弦表.阿耶波多是印度最早的数学家和天文学家,对他了解甚少,只知道他于499年在北印度恒河流域的比哈尔笈多王朝首府帕坦(Pātalipura)附近的古素玛(Kusumapura)写成这部著作.《阿耶波多历数书》由四部分123个命题组成,其中第二部分的33个命题是数学内容,但并没有详细的叙述,只给出了简要的说明.这也可能是由于当时为了便于记忆或是演讲稿的摘要.第二部分的命题12给出了正弦表的构造法,第一部分的命题10则给出了正弦差分表[15].

命题(Stanza)Ⅱ-12 正弦差的推导 第一个正弦值除以本身,然后减去所得商,就是第二个正弦差.同样,第一个正弦值减去后面所有正弦值和,除以第一个正弦值,就得其余的正弦差.

补遗 6.1	"正弦(sine)"词源
	英语中的"正弦(sine)"一词由梵文 *jyā-ardha*(半弦)演变而来,阿耶波多经常把它简写为 *jyā* 或 *jīyā*,后来当一些印度著作被译为阿拉伯文时,只是采用了该词的音译而非意译,即把它译为 jiba.由于在书面阿拉伯语中略去了元音,后人把辅音 *jb* 译为 *jaib*(胸膛).在12世纪,在一部阿拉伯三角学的拉丁文译本中,使用了与 *jaib* 对等的拉丁词 *sinus*,意思仍为胸膛,但引申意是港湾,英语中的 sin 即由 *sinus* 发展而来.

在印度三角学中,"第一正弦"s_1总是指$3\frac{3°}{4}=3°45'$的正弦.在半径为3438的圆中,若以分为单位,那么$s_1=225$.按本命题的规则,我们可以计算3°45′任意倍弧度的正弦,比如要计算s_2,即7°30′的正弦,从225中减去225得0(这一步中,第一和第二个正弦是相同的),再从225中减去$0+1=1$,即得第二个正弦差224,因此$s_2=225+224=449$.要计算s_3,从225中减去224得1,再用225除449得2,然后从225中减去$1+2=3$得222为下一个正弦差,于是11°15′的正弦$s_3=449+222=671$.一般地,第n个正弦s_n($n\times 3°45'$的正弦)可用下面的公式计算:

$$s_n = s_{n-1} + \left(s_1 - \frac{s_1 + s_2 + \cdots + s_{n-1}}{s_1} \right),$$

式中括号内的项表示第n个正弦差,其中的一些值列表如下:

命题Ⅰ-10 以分为单位的24个正弦差分别是225,224,222,219,215,210,205,199,191,183,

174,164,154,143,131,119,106,93,79,65,51,37,22,7.

上面的这些值与按前面的方法计算出的值略有差异,可能由于除法过程中不断地在正弦值间分配分数值的原因,也可能印度人最初并不是用这种方法计算这些正弦值,他们很可能使用了如希帕科斯的方法:90° 的正弦值等于半径 3438′,30° 的正弦等于半径的一半 1719′,45° 的正弦等于$\frac{3438}{\sqrt{2}} = 2431'$,对其它的角的正弦可应用毕达哥拉斯定理和半角公式求得.

只要按 3°45′ 的倍数作出 3°45′ 到 90° 间各角的正弦表,那么一差和二差的表也可由此作出.若印度人注意到了二差与正弦间的比例关系,给出命题 Ⅱ-12 中的规则也并不是困难的.随后几百年里,印度一些学者也作出了与此精确度差不多的正弦表.瓦拉哈米希拉(Varāhamihira,6 世纪)构作了半径是 120 的正弦函数和余弦函数表,并描述了这些函数间的标准关系.《苏利耶·悉檀多》(Sūrya-Siddhānta,也译为《太阳的知识》)是写于 7 世纪的一本天文著作,我们前面谈到的中国正切函数的计算很可能受该书的影响,这本书中还间接地提到了正割函数,但其中并未给出明确的函数表.书中第 3 章讲述晷影的诗文 21—22 中写道:"(求正午天顶距),先求正、余弦值,正弦、半径各乘晷测数,余弦除之即得正午影与弦."[16]

6.6.2 印度的估值技巧

有趣的是,在巴什迦罗(Bhāskara,12 世纪)之前,还没有一本天文著作中给出间距比 $3\frac{3°}{4}$ 更小的正弦表,而印度数学家却发展了这种估值方法,当然最简单的方法便是正弦表中的内插法.在此之前,7 世纪的波罗摩笈多(Brahmagupta,598)就已经用二阶差列出了更精确的内插表,用现在的符号表示,若 D_i 表示第 i 个正弦差(如阿耶波多命题 Ⅰ-10 中的数值),x_i 表示第 i 条弧,$h = 3\frac{3°}{4}$ 表示这些弧的间距,那么波罗摩笈多所给的计算结果是:

$$\sin(x_i + \theta) = \sin(x_i) + \frac{\theta}{2h}(D_i + D_{i+1}) - \frac{\theta^2}{2h^2}(D_i - D_{i+1}).$$

例如要计算 sin(20°),因为 $20 = 18\frac{3}{4} + 1\frac{1}{4}$,设 $18\frac{3}{4} = x_5$,代入上式得

$$\begin{aligned}\sin(20) &= \sin\left(18\frac{3}{4} + 1\frac{1}{4}\right)\\ &= \sin\left(18\frac{3}{4}\right) + \frac{1\frac{1}{4}}{2\left(3\frac{3}{4}\right)}(215 + 210) - \frac{\left(1\frac{1}{4}\right)^2}{2\left(3\frac{3}{4}\right)^2}(215 - 210)\\ &= 1105 + \frac{1}{6}(425) - \frac{1}{18}(5) = 1176,\end{aligned}$$

这里取近似整数,正弦值是相对于半径 3438 的圆而言.波罗摩笈多对此公式没有给出证明,他之后有印度数学家给出了合理的证明.另外借助 17 世纪欧洲人制定的标准的差分估值方法也可以说明上式的正确性.耐人寻味的是,波罗摩笈多也使用了代数公式去估算正弦值,而这种代数公式与他同时代的但比他稍早些的巴什迦罗(7 世初)已曾给出.巴什迦罗在 Mahābhāskariya 梵文诗中写道:

> 我简述不使用 225 和其它的正弦差来计算正弦的方法.从 180° 中减去(一个弧)度,以此弧度乘以所余弧,把所得值分别置于两个位置,在其中一个位置上用这个积去减 40 500,再取其四分之一,然后除以前一个位置的数后再乘半径,即可得(相对此半径的)正弦.[17]

巴什迦罗的公式用现在的符号表示,就是

$$R\sin\theta = \frac{R\theta(180-\theta)}{\frac{1}{4}(40\,500-\theta(180-\theta))} = \frac{4R\theta(180-\theta)}{40\,500-\theta(180-\theta)}.$$

如果用此式计算 $\theta = 20°$ 的正弦,得

$$3438\sin 20 = 3438\cdot\frac{4\cdot 20\cdot 160}{40\,500-20\cdot 160} = 1180,$$

这里取近似整数,误差为0.3%.

这里有两个问题出现,一、如何推导出这个代数公式?二、既然印度人已经用几何的方法给出了较精确的正弦表和标准的内插法,那为什么还要用代数的公式呢?对这两个问题,从古文献中几乎不能找到答案.所以我们用现代最简单的方法加以推测,思路如下,注意到当 $\theta = 0$ 和 180 时,正弦函数 $R\sin\theta$ 与抛物线函数 $P(\theta) = R\theta(180-\theta)/8100$ 的值都等于0,当 $\theta = 90$ 时都等于 R,因此他们可能注意到这对函数 $P\sin\theta$ 也是成立的.因为 $P(30) = (5/9)R$, $P(30)\sin(30) = (5/18)R$,他们接着应用下面的比例性质去调整 P 的值以得出当 $\theta = 30$ 时的正确值 $R/2$,

$$\frac{P-R\sin\theta}{P\sin\theta-R\sin\theta} = \frac{\frac{5}{9}R-\frac{1}{2}R}{\frac{5}{18}R-\frac{1}{2}R},$$

于是可导出等式

$$R\sin\theta = \frac{4P}{5-\frac{P}{R}},$$

由此可得出巴什迦罗的公式.[18]

这种导出估值公式的简单方法也出现在其它印度数学中:先给出合理的猜测,再通过修正使它与一些选定值的正确结果一致.但是并没有数学家给出他所谓"修正"的解释,因此很难得知他们是根据什么又是怎样得出结果的.事情也许很简单,数学家们为了锻炼他们的创造性思维而产生了这些聪明、漂亮的结果.在许多的天文计算中要用到正弦函数,因此在天文学家眼里,有一个正弦函数的准确估值是十分方便的,这使他们节省了在已发表的正弦表中去作反复的插值运算所花费的大量精力.正如我们在后面其它场合将会看到的那样,没有根据说印度数学家会拘泥于以某种特定的形式证明框架为基础的方法.因此,虽然他们通常也知道怎样去"证明"数学结论,但现存的文献表明,一旦他们确认了某一结论的合理性,就会让它一代代地传承下去.

6.6.3　阿耶波多和他的《阿耶波多历数书》

《阿耶波多历数书》(*Āryabhatīya*)数学卷中的内容除了三角学外,还有一些其它的内容,如不定方程的解法及步骤、一些问题的综述、数值计算、代数等.而阿耶波多(Ārhabhata)对其书中的这些方法并未给出解释.

第2卷中命题5给出了求立方根的方法.

命题Ⅱ-5　第二个非立方位(aghana)除以前一立方位立方根平方的3倍;第一个立方位减去商的平方与已得的立方根部分(purva)的三倍之乘积.(然后)从立方位减去商的立方.

文中的术语涉及已知数中数字的位置.从右向左数,第一、第四等位叫做立方位(ghana);第二、第五等位叫做第一非立方位(aghana);第三、第六等位叫做第二非立方位.但是正如计算12 997 875的立

方根的例子表明,某些步骤并不写出,很可能是受梵文诗体格式的限制.

	1 2 9 7 7 8 7 5	)2	第一个数 $2 \approx \sqrt[3]{12}$
	8		2^3
1 2	4 9	)3	$12 = 3 \times 2^2$
	3 6		$49 \div 12$ 取近似值 3,因为 4 太大
	1 3 7		$36 = 3 \times 2^2 \times 3$
	5 4		$54 = 3 \times 2 \times 3^2$
	8 3 7		
	2 7		3^3
1 5 8 7	8 1 0 8	)5	$1587 = 3 \times 23^2$
	7 9 3 5		$8108 \div 1587$ 取近似值 5
	1 7 3 7		$7935 = 3 \times 23^2 \times 5$
	1 7 2 5		$1725 = 3 \times 23 \times 5^2$
	1 2 5		
	1 2 5		5^3

上面算法的依据是 $(a+b)^3$ 的展开式,例如

$$23^3 = (20+3)^3 = 20^3 + 3 \times 20^2 \times 3 + 3 \times 20 \times 3^2 + 3^3.$$

命题 Ⅱ - 16 两影端距离与第一影长相乘,除以两影长之差,所得结果是 kotī(底线),这一线段与标杆长相乘,再除以第一影长,即得 bhujā(高,即杆长).

这一命题给出了通过测量杆顶到地面的不同影长求杆长的方法,如图 6.11 所示,DE 和 $D'E'$ 是两个高为 g 的杆,杆 AB 的投影 $DF = s_1$,$D'C = s_2$,影端距 FC 皆为已知,设杆 AB 长为 x,$AF = y$,把这一命题 2 表示为公式,即是

$$y = \frac{(s_2 + t)s_1}{s_2 - s_1} \quad \text{和} \quad x = \frac{yg}{s_1}.$$

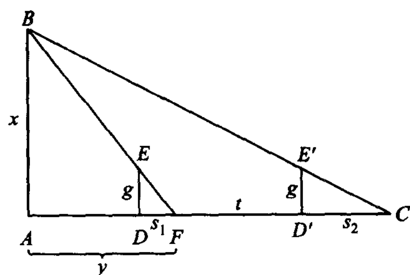


图 6.11 《阿耶波多历数书》中高度和长度的测量.

这一问题的形式和解决方法与中国的《海岛算经》中的问题 1 十分相似.

命题 Ⅱ - 19 项数减 1,然后除以 2,…,然后乘以公差,再加上首项后即得中项.再以中项乘以项数所得结果是(前面诸项的)之和.或首末项之和乘以项数的一半.

阿耶波多在此命题中给出了求首项为 a,公差为 d 的等差数列之和 S_n 的公式.这个公式相当于

$$S_n = n\left[\left(\frac{n-1}{2}\right)d + a\right] = \frac{n}{2}[a + (a + (n-1)d)]. \tag{6.3}$$

命题 Ⅱ - 20 数列之和乘以 8,再乘以公差,再加上首项的 2 倍减公差之差的平方,然后取平方根,再减去首项的 2 倍,再除以公差,最后加 1 后再除以 2,即可得数列的项数.

在与命题 Ⅱ - 19 同样的条件下,若已知 S_n,那么便可求出项数 n,公式如下:

$$n = \frac{1}{2}\left[\frac{\sqrt{8S_n d + (2a-d)^2} - 2a}{d} + 1\right]. \tag{6.4}$$

如果把式(6.3)改写为 n 的二次方程,即 $dn^2 + (2a-d)n - 2S_n = 0$,然后可得出(6.4)中 n 的值.虽然阿耶波多并未明确给出二次方程式,但一个世纪多之前的波罗摩笈多已给出了这样的公式,显然阿耶波多也知道这一公式.

命题 Ⅱ - 22 (自然)数列的平方和等于项数乘以项数加 1 之和、再乘以项数的 2 倍加 1,然后

除以 6.(自然)数列的立方和等于该数列和的平方.

这一命题给出了求前 n 个自然数的平方和 S_n^2 和立方和 S_n^3 的公式

$$S_n^2 = \frac{1}{6}n(n+1)(2n+1), \quad S_n^3 = (1+2+\cdots+n)^2.$$

实际上,阿基米德已经知道了前一个公式,而第二个公式几乎是显然的,至少当作一个假设,代入一些具体的数加以验证后即可归纳出来.我们并不能确切地知道阿耶波多怎样得出这些公式的.

《阿耶波多历数书》数学卷中最后的两部分是解一次整数不定方程的内容.我们将在波罗摩笈多著作的研究中详细讨论这样的问题.

6.7　印度对不定方程的研究

波罗摩笈多于 598 年生于印度南部,他一生的大部分时间在圭拉王国(Guyaras)的首都比拉马拉(Bhillamāla,现今拉贾斯坦邦的比马尔)度过,但在他生活的时代,比拉马拉已成为戒日王朝的领土.他本人的主要工作是他 30 岁时著述的《波罗摩修正历数书》(*Brāhmasphutasiddhānta*),他被大家称为 Bhillamālacarya—— 来自比拉马拉的老师.

与其它许多印度中世纪的数学著作一样,波罗摩笈多的数学著述也是作为天文著作的一部分出现,把数学方法应用于解决不同的天文问题.和阿耶波多一样,他的著作也是以诗文的形式表述,但他对问题解答的表述比其先人更详细,同时还给出一些实例加以说明.可能由于多年的撰抄或不要求记录解题各个步骤的口述传统,书中对许多方法的叙述与其所给实例的步骤不相符.我们所给出的现代的解释只是体现了他主要的思想.波罗摩笈多并不给出结论的证明,而只给出解决问题的算法.

6.7.1　一次(线性)同余

波罗摩笈多感兴趣的问题是,确定一整数,使得此整数被两个正整数除后所得余数为已知.用现在的术语来说,即他想确定整数 N,使得 $N \equiv a \pmod r$ 和 $N \equiv b \pmod s$,或者

$$N = a + rx = b + sy \quad 或 \quad a + rx = b + sy;$$

或者令 $c = a - b$,使得 $rx + c = sy$.

我们用波罗摩笈多著作中的如下例子叙述他的方法(也叫"库塔卡"方法):$N \equiv 10 \pmod{137}$,$N \equiv 0 \pmod{60}$,也可表示成一个方程 $137x + 10 = 60y$.

"库塔卡(Kuttaka)"方法:

> 用得出最小余数的除数除以得出最大余数的除数;余数分别相除;逐次求出商.

60 除 137 后,把余数依次相除.也就是说应用欧几里得算法直至得到非零余数.

$$\begin{aligned} 137 &= 2 \cdot 60 + 17, \\ 60 &= 3 \cdot 17 + 9, \\ 17 &= 1 \cdot 9 + 8, \\ 9 &= 1 \cdot 8 + 1. \end{aligned}$$

商的数列为

2
3
1
1

波罗摩笈多把 0 看作是第一个商,因为第一个除式是 $60 = 0 \cdot 137 + 60$.

> 最后的余数乘以一个假定的数,使得这个乘积加上或减去这个余数后可以被上一余数整除,然后可以求出这个乘数和商.

用某个数 v 乘以最后的余数是 1,使得 $1 \cdot v \pm 10$ 恰好可被上一余数整除,上面的例中即可被 8 整除.波罗摩笈多解释道,当商为偶数时用 +,奇数时用 -.此处因为 0 也是一个商,后一个方程变为 $1v - 10 = 8w$,取 $v = 18$、$w = 1$,可得出一列新的数

0
2
3
1
18
1

> 倒数第二项乘以它上面的一项,再加上其下面的一项[这一过程一直进行到这列数的最顶端],最顶端的数是[agrante].

18 乘 1 再加上 1 等于 19,然后用 19 替换"上面"的项,即 1,再去掉最后一项,连续进行这一过程直至只剩下两项.

0	0	0	0	0	130
2	2	2	2	297	297
3	3	3	130	130	
1	1	37	37		
1	19	19			
18	18				
1					

最顶端的项,(agrante) 是 130,因此 $x = 130, y = 297$ 是开始所求方程的一个解.但波罗摩笈多想得到更小的数值解,为此他先确定 N:

> 用得出最小余数的除数除以[The agrante];得到的余数乘以得出最大余数的除数,再加上较大的余数,得到的是除数乘积的余数.[19]

用 60 除 130 得到余数是 10,10 乘以 137 后再加上 10 得 1380,则有 $N \equiv 1380 \pmod{8220}$.然后波罗摩笈多通过用 60 除以 1380(因为 $N = 60y$)而解得 y 的值,同时计算得到一个新的 x 的值.因此 $y = 23$、$x = 10$ 是方程 $137x + 10 = 60y$ 的一个解.

虽然我们并不知道波罗摩笈多怎样给他的学生演示他的求解过程,但我们可以给出一个现代的解释.首先解方程 $137x + 10 = 60y$,按照欧几里得算法中出现的商逐步替换,上面的例子中共有 10 次替换.

$$137x + 10 = 60y, \qquad 60 = 0 \cdot 137 + 60,$$

$$
\begin{aligned}
&137x + 10 = (0 \cdot 137 + 60)y, && \\
&137(x - 0y) + 10 = 60y, && z = x - 0y, \\
&137z + 10 = 60y, && \\
&60y - 10 = 137z, && 137 = 2 \cdot 60 + 17, \\
&60y - 10 = (2 \cdot 60 + 17)z, && \\
&60(y - 2z) - 10 = 17z, && t = y - 2z, \\
&60t - 10 = 17z, && \\
&17z + 10 = 60t, && 60 = 3 \cdot 17 + 9, \\
&17z + 10 = (3 \cdot 17 + 9)t, && \\
&17(z - 3t) + 10 = 9t, && u = z - 3t, \\
&17u + 10 = 9t, && \\
&9t - 10 = 17u, && 17 = 1 \cdot 9 + 8, \\
&9t - 10 = (1 \cdot 9 + 8)u, && \\
&9(t - 1u) - 10 = 8u, && v = t - 1u, \\
&9v - 10 = 8u, && \\
&8u + 10 = 9v, && 9 = 1 \cdot 8 + 1, \\
&8u + 10 = (1 \cdot 8 + 1)v, && \\
&8(u - 1v) + 10 = 1v, && w = u - 1v, \\
&8w + 10 = 1v, && \\
&1v - 10 = 8w. &&
\end{aligned}
$$

波罗摩笈多已经提到了上面一系列方程中的最后一个，由这个方程通过观察即可得到：$v = 18$、$w = 1$. 其余的变量亦可逐次确定. 波罗摩笈多准确地给出了下面的过程.

$$
\begin{aligned}
u &= 1v + w = 1 \cdot 18 + 1 = 19, \\
t &= 1u + v = 1 \cdot 19 + 18 = 37, \\
z &= 3t + u = 3 \cdot 37 + 19 = 130, \\
y &= 2z + t = 2 \cdot 130 + 37 = 297, \\
x &= 0y + z = 0 \cdot 297 + 130 = 130.
\end{aligned}
$$

用现在的术语，可以说波罗摩笈多和秦九韶都对线性同余问题的解决感兴趣，但认真的分析表明，这两种方法是截然不同的，特别是印度数学家通常只研究含有两个同余式的组，而中国数学家研究含有更多同余式的组. 即使波罗摩笈多在研究与“中国剩余问题”类似的问题时，他也只计算其中的两个同余式. 如对“除 6 余 5；除 5 余 4；除 4 余 3；除 3 余 2，此数为何？”的研究，他先计算 $N \equiv 5 \pmod 6$ 和 $N \equiv 4 \pmod 5$，得到 $N \equiv 29 \pmod{30}$，然后再计算 $N \equiv 29 \pmod{30}$ 和 $N \equiv 3 \pmod 4$ 等等. 印度和中国对此问题研究的相似之处是他们都应用了欧几里得算法. 更有趣的问题是：这两个文明是否学习了希腊的算法？还是三个文明向另外更早的文明学来的？抑或这两个亚洲文明各自独立地发现了这一算法？[20]

波罗摩笈多和阿耶波多研究一次同余式的动因是由于天文学上的需要，这与中国研究同余问题的动因相同. 第 5、6 世纪的印度天文学深受希腊的影响，特别是天体的本轮观念，因此印度天文学家像希腊天文学家一样需要用三角学去计算天体的位置. 但印度天文学上还有一个重要的思想是大的天文周期，即所有天体（包括太阳和月亮）周而复始同处于零经度上的时间. 这一点颇似古代的中国，而在希腊人看来却并不重要. 当时认为在这个天文周期内所有的事件将重新发生. 阿耶

波多认为基本的周期是 Mahayuga(4 320 000 年)—— 它的最后四分之一(Kaliyuga) 始于公元前 3102 年.但波罗摩笈多却认为基本的周期是相当于 1000 Mahayugas 的 Kalpa.

无论如何,要计算天体的位置,就必须知道它们的平均运动.按经验去确定天体的运动是困难的,因此就要根据当前的观测值去计算.实际上,所有天体在运动周期的开始都约在同一个位置上.这样的计算就需要解一次同余式.

6.7.2 佩尔方程

印度数学家感兴趣的另一类不定式是形如 $Dx^2 \pm b^2 = y^2$ 的二次方程,它首先出现在波罗摩笈多的著作中.该方程的特例 $Dx^2 \pm 1 = y^2$ 现在一般称为佩尔方程(错误地以 17 世纪英国人 John Pell 的名字命名).有资料表明,古希腊人已经解出了这一方程的一些特例,但从现有的文献来看,印度人最早开始解决这一方程的一般形式.波罗摩笈多在研究"库塔卡" 方法时以实例的形式引入了解决这类方程的方法.

> 可以在一年内计算出(一个数)的平方乘以 92 再加 1 恰好等于一个平方数(的人便)可称为数学家.

下面来解这个方程 $92x^2 + 1 = y^2$,这里当然不需要一年的时间.波罗摩笈多的方法是这样开始的:

> [求出]一个根,由这个根的平方乘以乘数,再加上或减去一个假定的数.

先确定一个数,比如 1,如果 92 乘以 1^2,再加上 8(假定的数),那么和是一个平方数 100.于是满足方程 $Dx_0^2 + b_0 = y_0^2$ 的三个数 x_0, b_0, y_0 已经找到.为了便于叙述,我们记(x_0, y_0)为关于加数 b_0 的一个解,那么(1,10) 就是关于加数 8 的一个解,波罗摩笈多把这一解写成二行,

$$\begin{matrix} x_0 & y_0 & b_0 \\ x_0 & y_0 & b_0 \end{matrix}$$

或

$$\begin{matrix} 1 & 10 & 8 \\ 1 & 10 & 8 \end{matrix}$$

> 乘数乘以前一对数的积再加上后一对数的积等于后根.

也就是说,通过设 $y_1 = Dx_0^2 + y_0^2$ 就可求出"后根"y 的一个新值,本例中即

$$y_1 = 92(1)^2 + 10^2 = 192.$$

> 交叉相乘后求和得前根,加数等于加量或减量的积.[21]

当有一个新的加数 $b_1 = b_0^2$ 时,由 $x_1 = x_0y_0 + x_0y_0$ 或 $x_1 = 2x_0y_0$ 即可求出"前根"x_1 的一个新的值,也就是说$(x_1, y_1) = (20,192)$ 是关于加数 $b_1 = 64$ 的解,或 $92(20)^2 + 64 = 192^2$.这个结果可以直接证明,但事实上波罗摩笈多考虑了更一般的结论,如果(u_0, v_0)、(u_1, v_1) 分别是关于加数 c_0 和 c_1 的解,那么$(u_0v_1 + u_1v_0, Du_0u_1 + v_0v_1)$ 是关于加数 c_0c_1 的一个解.已知 $Du_0^2 + c_0 = v_0^2$ 和 $Du_1^2 + c_1 = v_1^2$,我们可用下面的等式验证上面的结果:$D(u_0v_1 + u_1v_0)^2 + c_0c_1 = (Du_0u_1 + v_0v_1)^2$.称这个新的解为$(u_0, v_0)$ 和(u_1, v_1) 的**复合**.波罗摩笈多归纳了他的基本方法:"这样得到的根除以最初的加数或减数后得到的是关于单位加数的根",在上例中,8 除以 20 和 192 后得到的$(\frac{5}{2}, 24)$是关于加数 1 的解.但是这些解中有非整数,这并不符合要求.于是他把这一结果合成为关于加数 1 的

整数解(120,1151),也就是 $92\cdot120^2+1=1151^2$.

作为解释波罗摩笈多方法的这个例子有一定的局限性.一般来说关于加数1的解是$\left(\frac{x_1}{b_0},\frac{y_1}{b_0}\right)$.但这并不能保证它们一定是整数或可以通过它们自己组合成为整数,波罗摩笈多只是给出了几个规则和例子,他并没有考虑整数解的存在性.他首先指出,只要已知关于加数1和另一加数的解,那么通过这种合成可以找到关于任何加数的其它解.一般地,一个已知的方程有无限多个解.

第二,如果已经找到关于加数4的一个解(u,v),那么可以找到关于加数1的一个解,即如果v是奇数,u是偶数,则

$$(u_1,v_1)=\left(u\left(\frac{v^2-1}{2}\right),v\left(\frac{v^2-3}{2}\right)\right)$$

便为要求的解.如果v是偶数,u是奇数,则

$$(u_1,v_1)=\left(\frac{2uv}{4},\frac{Du^2+v^2}{4}=\frac{2v^2-4}{4}\right)$$

是整数解.波罗摩笈多给出了前一种情况的例子:由$3u^2+4=v^2$的解$u=2$、$v=4$即可解出$3x^2+1=y^2$.

波罗摩笈多对加数4也给出了类似的法则,还给出了一些解某些特殊佩尔方程的法则.他的这些方法是正确的,但由于他并未给出证明,所以我们无从得知他是怎样发现这些方法的.印度数学家为何对这一问题十分感兴趣?这仍是个未解之谜.波罗摩笈多所给的一些例子中的变量x,y为天文学变量,在现实生活中并不能找到这样的事例.

不管怎样,佩尔方程已经成为印度随后几个世纪数学研究的一个传统问题,直到贾亚德瓦(Acarya Jayadeva,1000年)才给出了完整的解决,而中世纪最著名的印度数学家巴什迦罗Ⅱ(Bhāskara Ⅱ,1114—1185)给出的解法最易理解.

巴什迦罗在其几何著作《丽拉沃蒂》(Līlāvatī)中的主要目的是说明任何形如$Dx^2+1=y^2$的方程都有整数解.他先重述波罗摩笈多解这类方程的过程.他特别指出,一旦找到方程的一组解,那么通过作合成即可得到其它许多的解.更重要的是他讨论了所谓的循环法(chakravāla),其基本思想是,用"库塔卡"方法连续选择关于不同加数的数组解,最终可以得到关于加数1的解.这里我们以巴什迦罗的一个例子$67x^2+1=y^2$来说明他解一般形式$Dx^2+1=y^2$的方法.

> 选取前根和后根,确定被除数、加数和除数,然后找到乘数.

与前面一样,先选择关于任意加数b的一组解(u,v),本例中取(1,8)作为关于加数-3的一组解,接着对m解不定方程$um+v=bn$,这里$1m+8=-3n$;对任意整数t,可得$m=1+3t$,$n=-3-t$.

> 从已知系数中减去乘数的平方,或从(乘数的)平方中减去系数(使得余数最小),余数除最初的加数,得一新的加数;如果是系数减平方数,则取差的相反数.对应于乘数的商即是前根;由此也可得到后根.

就是说,选择一个数t,使得m^2尽可能地接近D,取$b_1=\pm\frac{D-m^2}{b}$(可能为负)作为新的加数,那么得到的前根$u_1=\frac{um+v}{b}$,后根$v_1=\sqrt{Du_1^2+b_1}$.本例中,巴什迦罗为使m^2接近67,他取$t=2,m=7$,那么$(D-m^2)/b=(67-49)/(-3)=-6$.由于是系数减平方数,所以新的加数是6.那么前根$u_1=\frac{1\cdot7+8}{-3}=-5$,由于这些根都是平方数,所以$u_1$取正.同时,$v_1=\sqrt{67.25+6}=$

$\sqrt{1681} = 41$,即(5,41)是关于加数6的一组解.

接着,把得到的解置于一边进行重复运算,数学家们称这种方法为"循环法",这样可以找到关于加数[或减数]4,2或1的整数根;已知了关于加数4和2的根,通过合成便可以得出关于加数1的根.[22]

人物小传	巴什迦罗(1114—1185)(Bhāskara)
	巴什迦罗(有时也称巴什迦罗 Ⅱ,因为此前还有一位同名数学家)生于印度南部的一个书香门弟之家,他的父亲是一位有名学者和天文学家.巴什迦罗曾长期出任乌贾因天文台的台长,他以在数学、天文学和机械工艺方面的出众才能为后人倍加称道.他的孙子在1206年受国王特许建立了一个学院来研究巴什迦罗的著作.与其先人一样,他的主要著作《天文系统极致》(*Sinddhāntaśiromani*)也是研究天文学知识的,巴什迦罗的两本数学著作分别是研究算术的《丽拉沃蒂》(Līlāvatī,女子名,意思是"漂亮的"或"可爱的")和研究代数的《算法本源》(*Bījaganita*).

巴什迦罗指出,如果上面的运算可以重复进行,那么最终可以得到关于加数4,2的解,关于1的解也可得出.前面已经提到,由关于加数或减数4的解可求得关于加数1的解,同样也可容易地求出关于加数或减数2以及关于减数1的解.在继续分析这一例子前,我们需要讨论两个问题(并不是巴什迦罗提出的).第一,为什么此方法的每一步都给出整数解?第二,为什么重复这一方法可以得到关于加数 ±4、±2 和 ±1 的解?

下面回答第一个问题,注意到通过把关于加数 b 的第一组解(u,v)和关于加数(m^2-D)的解$(1,m)$相结合可以得到巴什迦罗的方法,于是$(u',v')=(mu+v,Du+mv)$是关于加数$b(m^2-D)$的一组解.在这一结果中除以 b^2 可得到关于加数$\frac{m^2-D}{b}$的解$(u_1,v_1)=\left(\frac{mu+v}{b},\frac{Du+mv}{b}\right)$.因此显然 m 必须满足 $mu+v$ 是 b 的整数倍数.容易证明,如果$\frac{mu+v}{b}$是整数,那么

$$\frac{m^2-D}{b} \text{和} \frac{Du+mv}{b}=\pm\sqrt{Du_1^2+b_1}$$

也是整数,虽然巴什迦罗的著作中并没有给出证明.[23]

第二个问题的答案是,因为必须选择的使 m^2-D"最小".但是,最终能得到关于加数1的解的证明是十分困难的,直到1929年才有这方面的证明发表.[24] 可以这样说,巴什迦罗和贾亚德瓦(Jayadeva)虽给出了求解过程,但并没有给出证明,他们给出了一系列的例子来说明其方法的正确性.事实上,可以看出,"循环法"可以给出二次不定方程的最小可能解,由此即可求得其它的解.[25]

我们继续分析巴什迦罗的例子,由 $67\cdot1-3=8^2$ 可得 $67\cdot5^2+6=41^2$,再解能使$|m^2-67|$最小的方程 $5m+41=6n$.合适的选择是 $m=5$.那么$(u_2,v_2)=(11,90)$是关于加数 -7 的解,即 $67\cdot11^2-7=90^2$.接着再解方程 $11m+90=-7n$.取 $m=9$ 得到$(u_3,v_3)=(27,221)$是关于加数 -2 的解,即 $67\cdot27^2-2=221^2$.已经得到了关于 -2 的解,将(27,221)作变化即可得到$(u_4,v_4)=(11\,934,97\,684)$是关于加数4的解,然后巴什迦罗再把$(u_4,v_4)$除2,便得到方程 $67x^2+1=y^2$ 的解 $x=5967,y=48\,842$.

6.8 代数与组合学

佩尔方程的解法可以看作是中世纪印度数学发展的高峰.除此之外,波罗摩笈多和巴什迦罗的著作中还有其它的数学内容,如正负数的计算、解二次方程、线性方程组和组合学的基本结果.

6.8.1 代数技巧

中世纪的印度数学家多数是代数学家,他们的著作中有许多关于正负数、分数及代数表达式的计算法则,他们解释了怎样解一个或多个变量的线性或二次方程,给出了算术和几何级数的求和、平方数、立方数和三角形数,不尽根的计算,怎样计算排列和组合数.这些法则都是正确的,但他们多数并没有说明怎样得到这些法则并且这些法则为什么是正确的.

现存的文献中也有一些关于面积和体积的计算公式.从现代的角度看,他们在这方面并未取得很大的成就,严格意义上讲这些公式并不完全正确,许多只是取近似值,但他们又并没有说明这些只是近似的计算公式.这些公式与书中其它那些严格成立的法则一样好像是绝对正确的.当然在大多数情况下,也是既没有证明也没有讨论其所得结果的正确性.

一般认为,很可能印度数学家对希腊式的证明没兴趣,虽然他们发现一些规则且得出许多结果,但并没有写出得到它们的推导细节的想法.实际上,有一种说法是,特别是对一些优秀的成果,不给出推导细节的原因是防止成果被剽窃,就是说,如果一个数学家想把某一法则说成是自己的成果,但他又不能给出推导过程,那么他将受到追究.[26]

虽然没有书面推导过程,印度数学家们必须说明他们结果的正确性,他们的学生也一定会询问为什么这些法则和方法是正确的,因此他们必须给出解释和回答,遗憾的是他们的解释(即证明和推理)并没有被记录下来.

下面我们将讨论印度的一些数学法则,首先看基本的算术法则.例如,波罗摩笈多给出的正负数的运算法则:

> 两正量的和是正的,两负量的和是负的;一个正量和一个负量的和是它们的差,如果它们(绝对值)相等,则其和为零…….正量减正量,负量减负量就从(绝对值)大的减去小的,如果从小的减去大的,其差就反号…….从负量中减去正量,从正量减负量,它们必须合并.负量与正量的积得负,两个负量的积为正,两正量的积为正…….正量除以正量或负量除以负量得正量…….负量除以正量或正量除以负量得负量.[27]

在乘法和除法法则中,零和其它数一样参与运算.波罗摩笈多指出,零乘以任何量都得零,但却宣称零除以零也等于零,并认为"零除正量或负量等于以零作分母的一分数."[28] 巴什迦罗进一步解释道,"以零为分母的量称作无穷量,这个以零作分母构成的量,无论加入或取出多少量,都不会发生任何变化;就像永恒的上帝,历经宇宙洪荒或创生时期而没有任何变化一样,虽然那时有大量物种灭亡或产生."[29] 然而,当需要时可以进行零的乘除运算,因为虽然"零[与一个确定量的]积等于零,……,但如果下面还有其它的运算,那就必须保留这个积作为零的倍数…… 如果在后面的运算中有零是除数时,那么这个确定的量保持不变."[30] 于是巴什迦罗提出这样的问题,"乘以0后再加上此乘积的一半,再乘3除0等于63的数是多少?"[31] 他实际上考虑了如下方程:

$$\frac{3\left(0x+\frac{1}{2}0x\right)}{0}=63,$$

由于分式中分子中的 0 可以提取公因子然后“约掉”，所以方程变为 $3x + \frac{3}{2}x = 63$，它的解是 14.

巴什迦罗也清楚地认识到了分配律：

> 被乘数 135，乘数 12，积（乘数相继乘以被乘数的各个数字）—1620，或把乘数分解为 8 和 4 之和，然后这两个数分别乘以被乘数，再把所得两积相加，同样得到 1620……，或把 12 分为两个数字 1 和 2，然后这两个数字分别乘以乘数，再根据数字的位置把两积相加，亦得 1620.[32]

波罗摩笈多还以同样的形式给出了二次求根公式：

> 将绝对数置于平方项及简单未知项的另一端. 绝对数乘以平方项[系数]的 4 倍，加上简单未知项[系数]的平方. [其和]的平方根减去简单未知项[的系数]，把结果除以平方项[系数]的 2 倍，得未知数[值].

他以解方程 $x^2 - 10x = -9$ 为例加以说明：

> 绝对数[-9]乘以平方项[系数 -36]的 4 倍，加上简单项[系数]的平方[100]，(得 64)，平方根[8]减去简单项[系数 -10]，把余数除以平方项[系数]的 2 倍[2]，得未知数的值是 9.[33]

实际上这个方程还有另一个正根，即当平方根小于一次项系数时取平方根的相反数时得到的根. 虽然波罗摩笈多在这里没有提及另外的这个根，他的这段话可以改译成方程 $ax^2 + bx = c$ 的求根公式：

$$x = \frac{\sqrt{4ac + b^2} - b}{2a}.$$

巴什迦罗还研究了多个根的情况. 他解二次方程的基本技巧是利用配方，即在方程 $ax^2 + bx = c$ 的两端加上适当的数，使左端成为一个完全平方式：$(rx - s)^2 = d$，然后解 x 的方程 $rx - s = \sqrt{d}$. 同时他指出，“如果方程绝对边的根的值小于包含在有未知数那边的根中的带负号的数(s)，那么令它为正或负，即可找到未知数的两个根.”[34] 也就是说，如果 $\sqrt{d} < s$，那么 x 有两个根，分别是

$$\frac{s + \sqrt{d}}{r} \quad 和 \quad \frac{s - \sqrt{d}}{r}.$$

但巴什迦罗又自拆篱墙，因为他说“这只是在某些情况下成立”. 下面我们以两个例子解释他的意思.

> 一群顽猴不知数，八分之一的平方钻进林中争攀树，一十二只在山间，狂奔乱跑穷追逐，回声惹得群猴怒. 这群猴共有多少个？

巴什迦罗把它写成方程 $\left(\frac{1}{8}x\right)^2 + 12 = x$，方程两边同乘 64 化简后得 $x^2 - 64x = -768$，两边再加上 32^2 后得 $x^2 - 64x + 1024 = 256$，取平方根 $x - 32 = 16$，然后他写道，“在绝对边的数小于有未知数那边根中的带有负号的已知数.” 因此 16 前可以取为正或负，最后他总结道，“可以得到未知数的两个根 48 和 16.”

> 一群顽猴不知数，五分之一减三的平方去钻洞，另见一只爬上树. 这群猴共有多少个？

方程变为 $x^2 - 55x = -250$，巴什迦罗求出了两个根 50 和 5，他说“但不能取第二个[根]，因为它是不恰当的，人们并不满意负的绝对数.”[35] 这里的负数并不是来自方程本身，而是来自于问题本身

人们不能从5个猴子的$\frac{1}{5}$中减去3个.对于我们现在来看应该有一正一负两个根的二次方程,巴什迦罗只是给出一个正根.他从来没有给出有两个负根或实根不存在的二次方程,也没有给出有无理数根的二次方程,他所给的每个例子中的平方根总是有理数.

印度数学家也掌握了含有多个未知数的方程.印度南部迈索尔(Mysore)的数学家马哈维拉(Mahāvīra,9世纪)在他著名的《计算方法纲要》(*Ganitasārasaṅgraha*)中提出了与百鸡问题类似的一个问题,"鸽5值3钱,鹤7值5钱,鹅9值7钱,雀3值9钱,为讨国王欢喜,某人欲百钱买百禽,问禽、钱各几何?"[36]

为解此问题,马哈维拉给出了一个很复杂的算法.另一方面,巴什迦罗提出了同样的问题,而他给出的解法清楚地说明了为什么这个问题有多解的原因.他设鸽、鹤、鹅、雀数分别为未知数,这里我们记为d、c、s和p,根据禽的数目和价钱得到两个等式,

$$3d + 5c + 7s + 9p = 100,$$
$$5d + 7c + 9s + 3p = 100.$$

消去d,得一方程$c = 50 - 2s - 9p$,取p的任一值如4,得到一个标准的不定方程$c + 2s = 14$,它的解是$s = t, c = 14 - 2t$,其中t为任意值,由此可得$d = t - 2$.再令$t = 3$,则得$d = 1, c = 8, s = 3, p = 4$,因此鸽、鹤、鹅、雀数分别是5、56、27、12,值钱分别是3、40、21、36.他还指出,如果选不同的t的值,将得到不同的答案,"通过取不同的假定值,便可得出许多答案."[37]

6.8.2　组合学

虽然印度给出的组合规则没有证明,但它们是世界上组合规则的最早记录,比如苏斯鲁塔(Susruta)的医学论文(约写于公元前6世纪)中给出了从六种不同的味——苦、酸、咸、涩、甜和辣中分别取1、2、3、4、5、6种的组合方法有63种.[38]就是说,六种味中取2种的组合数是15,3种的组合数是20,等等.同时代一些其它的著作也包含有与组合相关的内容,但这些例子中的数取得都很小,简单的枚举就能给出答案,而我们并不知道他们是否已得到完整的组合公式.

6世纪的瓦哈米希拉的著作中给出了一些稍大的组合数,"如果16个一组的物体以4种不同方式变化,那么结果是1820"[39],就是说瓦哈米希拉想从16种成份中取4种组成一种香料,他计算出共有1820($= C_4^{16}$)种可能的方法.我们猜测他知道计算组合数的方法,因为他似乎不大可能去一一枚举这1820种组合.当时的印度文献中并没有求组合数的公式,但瓦哈米希拉的著作中好像提示了一个类似于推导帕斯卡三角的标准方法,可以逐一地得出这些组合数.[41]

补遗 6.2

印度人研究数学的原因

印度人为什么对数学那么感兴趣?我们可以从他们的数学著作所涉及的问题的类型得到一些答案,尽管与中国人有所不同,问题并不是完全的"实用型".在马哈维拉的《计算方法纲要》(*Ganitasārasaṅgraha*)一书的引言中可以看到对这一问题的更一般的答案:

在那些世界性的事务中,如在吠陀教……等宗教事务中总要用到计算.在有关情感、财富分配、音乐话剧、烹饪艺术、医疗、建筑、韵律学、诗歌、逻辑学、语法学等学科中,计算的科学都受到高度重视.在涉及太阳和其它天体的运行、日月食和行星连珠等问题时数学也是十分有用的.有关计数,海岛、海洋、山脉的径周,大范围居民区的规划,和居民居所的设计等需要用到计算.[40]

但在9世纪,马哈维拉给出了计算组合数的详细算法:

> 求从已知数目的物体中取不同数目的组合数法则:从1开始,逐次加1,直至加到与已知的数目相等,并按正常的顺序写下,再按逆序分别写成上下两行,如果上面行中的数1,2,3,等等从右向左取的积,被下面行中的数1,2,3等等亦从右向左取对应的乘积除,这样可以求出各种情况下组合数的结果.[42]

这一算法可以表述成现代的公式

$$C_r^n = \frac{n(n-1)(n-2)\cdots(n+r-1)}{r!}.$$

但他并没有给出这一算法的任何证明.他把这一法则应用到两个问题上,一个是像前人一样计算调味品的组合数,另一个是项链上宝石的组合数,这些宝石可能是钻石、蓝宝石、翡翠、珊瑚和珍珠.

巴什迦罗在马哈维拉的基础上进一步指出,"这是一个普遍成立的法则,它在诗歌中可用于找出音步的变化;在建筑艺术中计算开门的变化;……,在医学中计算不同调味的组合."[43] 作为这种计算的一个例子,巴什迦罗写道,"一个技术高明的工程师为贵族设计了一座带八个门的宏伟宽敞的大厦,说出开1扇门,开2扇门,开3扇门等等的组合数."然后他计算出有8种方法打开1个门,28种方法打开两个门,等等.他断言,打开门的总方法数是225(不包括所有的门都关着的情况).

巴什迦罗还计算了 n 个元素的排列数是 $n!$,他提出并解答了下面的问题.

> 散卜神(Sambhu)的十只手中拿十件东西,绳、象牙、蛇、鼓、头盖骨、三叉戟、床架、匕首、弓、箭,若十只手交换拿这十样东西,共有多少种不同拿法?[44]

尽管中世纪印度数学家主要研究代数问题,但由于天文学的需要,他们对三角学也十分熟悉.在印度南部喀拉拉地区,14世纪到16世纪印度数学家发展了用无穷级数表示正弦函数(或相关函数)的方法,但后来却中断了,而这种方法在欧洲却是在17世纪才开始出现.

实际上,喀拉拉学派开始对问题的结论给出证明与推导,这是印度人给出数学证明与推导的惟一时代.因为这些结果与微积分有密切关系,所以我们把它们放在第12章加以讨论.

6.9 印度 - 阿拉伯十进位值制数系

作为本章的结束,我们将简要地讨论一下我们今天所使用的十进位值制数系的起源.人们通常称这种数系为印度 - 阿拉伯数系,这是因为这种数系最早可能源于印度,后来通过阿拉伯传到了西方.数字1到9、位值的概念和0的使用是这个数系最重要的组成.这里我们将要概括介绍对这三种思想的起源和发展的最近的研究.

表示数字1到9的符号在印度的婆罗门教文献中已经出现,它们至少可追溯到阿育王时代(公元前3世纪中叶),许多刻在柱子上的国王的法令中有这些数的符号.约在8世纪,伊斯兰国家入侵印度南部,同时征服了地中海地区大部分的国家,然后他们采用了这些数字.一个世纪后,这些数字在西班牙出现,再晚些又在意大利和欧洲出现.

比数的符号形式更重要的是数的位值制,对这方面历史的研究还很薄弱.巴比伦有位值制,但却是60进制的.虽然这一位值制被希腊人基于天文的需要而采用,但总体来说对他们的书写数系却没有什么影响.中国古代早已有了以十为基数的乘法累数制数系,这很可能源自中国的计算板(筹算盘),它由许多列组成,每列表示10的不同幂.在印度,虽然有从1到9的数字符号,同时总有

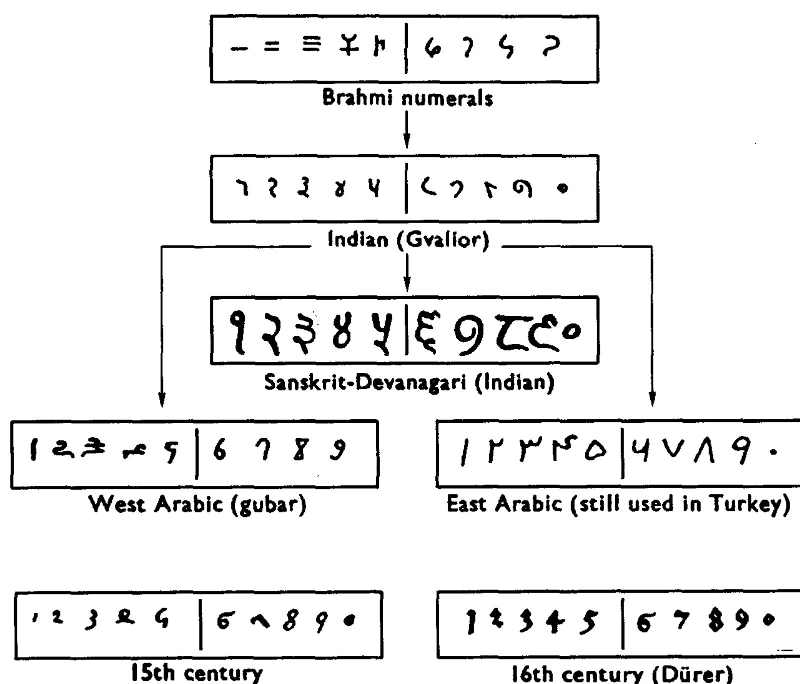


图6.12　现代数码的发展演化.(资料来源:《数字和数的符号》(*Number Words and Number Symbols, A Cultural History of Numbers*),Karl Menninger 著,Paul Broneer 译自德文修订版,英译版权 ©1969, The Massachusetts Institute of Technology.)

表示10到90的符号,对较大的数的表示与中文书写形式一样,是由前9个数字之一与表示100或1000的符号组合而成.古代印度和中国一样,也较早就使用了乘法累数制.在阿耶波多的著作中就列出了10的不同幂的名称:"*dasa*[十], *sata*[百], *sahasra*[千], *ayuta*[万], *niyuta*[十万],……[45].

约公元600年,印度人显然放弃了表示大于9的数的专门符号,而开始以今天熟悉的位值形式来使用他们的九个符号1,…,9.但这种用法最早并不源于印度.约662年叙利亚的牧师塞弗鲁斯·塞博克斯(Severus Sebokht)的著作残本中有一段评论说,印度人使用一种"用9个符号做计算"[46]的方法.他只提到9个符号,并未提到符号0.但在巴克沙利手稿(1881年在印度南部巴克沙利村庄发现的数学手稿)中,使用了位值制并用一个点表示0.有很好的证据证明这是7世纪的手稿.可能塞弗鲁斯没有把点看作是一个"符号".在同一时期的其它印度数学著作中,为了适合诗体文本的格式,通常将数写成半位值制,例如在马哈维拉的著作中,用特定的词代表数:月(moon)代表1,眼(eye)代表2,火(fire)代表3,天(sky)代表0,那么词火-天-月-眼(fire-sky-moon-eye)就表示2103,月-眼-天-火(moon-eye-sky-fire)代表3021,注意到位置是从右向左的.

奇怪的是,第一批用十进位值制数码标记年代的碑文是在柬埔寨发现的,最早的一块出现在683年,其中萨卡(saka)纪元第605年用三个数字表示,中间的0用一个点表示,而第608年的三个数字中间一个是用与我们现在的0相同的符号表示.在中国工作的印度学者在718年为中国编写的天文著作《九执历》中,也用点表示0作为十进制的一个符号.虽然作者并没有给出其它数字的印度符

号,但他给出了位值制的原理:"用[印度]数码进行乘除运算,每个数码用一笔写成.当计到10时,就要进高一位.在每个空位上放一个点,于是每一位上都记有一个数字.这样就不会发生定位错误.有了这种数码,运算就变得容易多了……"[47]

为什么印度7世纪初就提出并引入了包括0的符号的位值制?我们仍不能给出这个问题确定的答案.但有些学者认为印度数码的位值制很可能是受中国算板的影响而产生的.中国算板是可以随身携带的,很可能中国商人把这种算板带到印度.实际上,东南亚地处印度和中国文化的影响的碰撞点,那里必定有广泛的文化交流.很可能是印度人受到了只用9个符号思想的影响,但他们仍使用他们已经用惯了的符号,并进一步改进了中国的算筹数系.他们不是根据数的位置交替使用两种符号,而是不论什么位置,同样的数字都用同样的符号.因为他们不只要把数列在算板上,同时还需要把数写成某种书面形式,所以不得不用点或后来的圈表示算板上的空位.[48] 如果这种说法正确,那么有讽刺意味的是,印度科学家在8世纪初又将这种新的数系重新引入了中国,作为对他们的回报.

不管怎样,我们可以断言,完整的整数位值制8世纪时在印度已经存在了,即使是将最早刻有十进位值制数字的石碑年代确定在公元870年.在此之前,这种位值制早已传播到了中国和伊斯兰文化的中心巴格达.但要指出的是,虽然中国人已经在算板上使用了十进制分数,但在印度却从未有过这方面的记录.是伊斯兰人引入了分数的十进制从而完善了印度的十进位值制.

习　题

中国测量问题

1.《海岛算经》中的问题3:测量方城 $ABCD$,在距 E、F 点10尺处立两杆(图6.13),从点 E 移动5尺到点 G,从点 D 发出的视线 DG 与 F 交于点 H,且 $HE = 3\frac{93}{120}$ 尺.视点再移动到点 K,使得 $KE = 13\frac{1}{3}$ 尺,线 DK 过点 F.求 DC 和 EC.[49](刘徽求得 $DC = 943\frac{3}{4}$ 尺,$EC = 1245$ 尺.

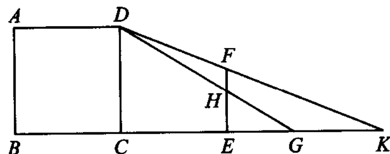


图 6.13 《海岛算经》问题3.

2.《九章算术》第九章第24题(这是一个基本问题的典型例子,刘徽正是在此例的启发下写了《海岛算经》):今有井,径5尺,不知其深.立5尺木于井上.从木末望水岸,入径4寸.问:井深几何?[50]

中国的不定分析问题

3.求解《数书九章》问题 Ⅰ - 1:$N \equiv 1 \pmod 1$,$N \equiv 1 \pmod 2$,$N \equiv 3 \pmod 3$,$N \equiv 1 \pmod 4$.

4.求解《数书九章》问题 Ⅰ - 5:$N \equiv 32 \pmod{83}$,$N \equiv 70 \pmod{110}$,$N \equiv 30 \pmod{135}$.

5.《张邱建算经》中的问题*:今有清酒一斗直7钱,普通酒一斗直3钱,劣酒三斗直1钱.若10钱买酒十斗,问每种

* 这一题与原著中的不一致.

酒及直钱分别是多少?

解数值方程问题

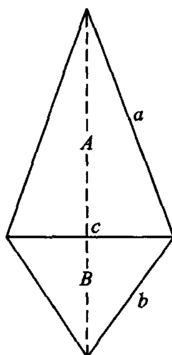


图 6.14 《数书九章》中的面积问题.

6. 秦九韶《数书九章》中关于求面积的数值方程. 如图 6.14 所示, 下面三角形的面积等于 $B = \frac{c}{2}\sqrt{b^2 - \left(\frac{c}{2}\right)^2}$, 上面三角形的面积等于 $A = \frac{c}{2}\sqrt{a^2 - \left(\frac{c}{2}\right)^2}$. 那么整个图形的面积 $x = A + B$. 证明 x 满足 4 次方程 $-x^4 + 2(A^2 + B^2)x^2 - (A^2 - B^2)^2 = 0$. 如果设 $a = 39, b = 20, c = 30$, 那么这就是秦九韶著作中所解的一个方程.
7. 用秦九韶的数值解法解下面两个方程. 它们是秦原著中的方程:
   (a) $16x^2 + 192x - 1863.2 = 0$
   (b) $-x^4 + 15\,245x^2 - 6\,262\,506.25 = 0$

   用秦九韶的数值方法解《杨辉算法》中的四个问题:
8. 矩形面积 864, 宽比长少 12, 求长与宽.
9. 矩形面积 864, 长与宽的差是 12, 求长与宽之和. 在杨辉的图中(图 6.15), 四周是四个相等的矩形, 中间是一个边长等于外围矩形长宽差的正方形.
10. 直田积八百六十四步, 只云长、阔共六十步. 欲先求长步, 问得几何. (与 9 同图)
11. 矩形面积是 864, 长、2 倍宽、3 倍长宽和与 4 倍长宽差之和等于 312, 求矩形宽.
12. 用秦九韶的方法解简单二次方程 $x^2 = 55\,225$. 并与第一章中的这个方程的解法作比较.
13. 用秦九韶的方法解简单四次方程 $x^4 = 279\,841$. 说明在解的过程中, 帕斯卡三角的四阶系数 4 6 4 1 怎样出现.
14. 朱世杰的一个问题: 圆周长的平方根与圆面积之和等于 114. 求圆周长与半径. (朱用一个四次方程解此问题.)

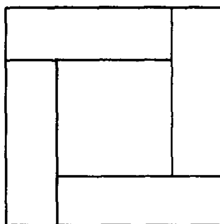


图 6.15 《杨辉算法》中的图.

印度三角学问题

15. 用阿耶波多的方法计算第 4、5、6 个正弦差, 然后与命题 Ⅰ - 10 中的结果比较. 再用这些值计算 15°、18°45′ 和 22°30′ 的正弦值(半径取 3438). 然后与用半角公式计算出的结果作比较.
16. 用图表计算器或微积分技巧证明用巴什迦罗 Ⅰ 的代数公式估算 0 到 180° 正弦值的误差不大于 1%, 求其中误差最大的一些值.
17. 用阿耶波多的方法计算 13,312,053 的立方根.

印度的不定分析问题

18. 波罗摩笈多论同余著作中的问题: 已知太阳在 10,960 天内沿黄道运行 30 圈. 当太阳运行了整数圈加上 8080/10,960 圈, 即"当太阳运动的余数是 8080," 它(离最初出发点)经过了多少天? 如果 y 是所求天数, x 是运动圈数, 那么由于太阳转 30 圈要 10,960 天, 则 x 圈要 $(1096/3)x$ 天. 因此, $y = (x + 808/1096)(1096/3)$, 或 $1096x + 808 = 3y$, 接着解 $N \equiv 808 \pmod{1096}$, $N \equiv 0 \pmod{3}$.
19. 解同余式 $N \equiv 23 \pmod{137}$, $N \equiv 0 \pmod{60}$.
20. 用波罗摩笈多的方法解不定方程 $1096x + 1 = 3y$. 已知这个方程的一个解(关于加数 1), 则容易通过乘法求出加数为其它值的方程的根, 如 $1096x + 10 = 3y$.
21. 证明波罗摩笈多的"库塔卡"法则. 首先注意到, 由欧几里得算法, 可以把两个正整数的最大公因子表示成这两个数的线性组合. 其次, 这个算法存在的条件是最大公因子必须整除"加数". 波罗摩笈多并未提出这一点, 但巴什迦罗和其他数学家确已论及.

22. 分别用印度和中国算法解同余式组 $N \equiv 5 \pmod 6 \equiv 4 \pmod 5 \equiv 3 \pmod 4 \equiv 2 \pmod 3$. 比较这两种方法的异同.
23. 用中国算法解同余式 $N \equiv 10 \pmod{137} \equiv 0 \pmod{60}$,并把解的过程与用“库塔卡”方法求解的过程作比较.
24. 根据欧几里得算法分别用中国和印度的方法解不定方程 $17n - 1 = 75m$,并把结果作比较.
25. 已知 $Du_0^2 + c_0 = v_0^2$ 和 $Du_1^2 + c_1 + v_1^2$,试证,$D(u_0v_1 + u_1v_0)^2 + c_0c_1 = (Du_0u_1 + v_0v_1)^2$.
26. 用波罗摩笈多的方法解方程 $83x^2 + 1 = y^2$,已知(1,9)是关于减数2的一组解.
27. 如果(u,v)是 $Dx^2 - 4 = y^2$ 的一组解,那么

$$(u_1, v_1) = \left(\frac{1}{2}uv(v^2+1)(v^2+3), (v^2+2)\left[\frac{1}{2}(v^2+1)(v^2+3) - 1\right]\right)$$

是 $Dx^2 + 1 = y^2$ 的一组解,无论 u, v 是奇数或偶数,u_1, v_1 均为整数.
28. 已知(1,3)是关于减数4的一组解,用习题27的方法解 $13x^2 + 1 = y^2$.
29. 如果(u,v)是 $Dx^2 + 2 = y^2$ 的一组解,试证$(u_1, v_1) = (uv, v^2 - 1)$是$Dx^2 + 1 = y^2$ 的一组解. 如果(u,v)是 $Dx^2 - 2 = y^2$ 的一组解,推导类似的公式.
30. 用巴什迦罗“循环法”解方程 $6x^2 + 1 = y^2$. 答案是 $x = 226\,153\,980, y = 1\,766\,319\,049$.

印度的几何问题

31. 解马哈维拉的下面问题:“春天的晚上,…… 花园里,月光融融,垂柳拂地,花香弥漫,鸟儿轻吟,蜂儿淘醉在花蜜的世界里 …….寓所里,年轻女士幸福地偎依在丈夫身旁.一不小心,丈夫扯断了妻子的项链,链上的珍珠脱落,三分之一滚到了女佣那边,六分之一落到床上,剩下的珠子中的一半(再剩下的一半,…… 共计6次)溅到了各个角落,最后还剩下1161颗珠子,…… 求项链上原有的珍珠数.”[51]
32. 马哈维拉的另一问题:有四只水管向井中注水,如果只用一只,分别需要1/2,1/3,1/4和1/5天注满水井,当四个水管同时向井中注水时,问:要多长时间注满水井?且每个水管各注人多少水.
33. 巴克沙利手稿中的问题:甲一日可行5里(yojanas),当他走了7天后,乙以每日9里的速度追赶,问:几日后乙可赶上甲?
34. 马哈维拉的问题:3只孔雀值2钱,4只鸽子值3钱,5只鹅值4钱,6只鹌鹑值5钱.今欲以56钱买72禽,问:每种禽各多少只?
35. 马哈维拉的另一问题:两行人甲和乙捡到一钱包,甲对乙说:“如果我得到钱包里钱的一半,那么我将和你有相同的钱”.乙对甲说:“如果我得到包内钱的三分之二,再加上我手里的钱,那么我的钱将是你现有钱的三倍”.问:甲、乙各有钱多少,包内钱是多少?

讨　论

36. 设计一节数论课,用秦九韶的方法演示中国剩余定理.
37. 设计一节学微积分之前的课,演示用综合的方法求多项式方程的数值解.
38. 比较刘徽和阿耶波多测算远处物体距离和高的方法.
39. 设计一节数论课,用波罗摩笈多的方法求解形如 $rx + c = sy$ 的不定方程.
40. 在印度人眼中,为什么他们认为用代数方法估算正弦函数的值要优于几何或插值的方法?
41. 比较古代中国和印度的数学观.
42. 参考16世纪访问中国的耶稣会士利玛窦(Matteo Ricci)的故事,写一篇关于他对中国数学研究影响的报告.

文献和注解

正如第一章中的参考与注释所说,对中国数学最好的综述是李约瑟(J. Needham)著的《中国科学技术史》(卷3)(*Science and Civilization in China*, vol. 3, Cambridge University Press, 1959)和李俨、杜石然著的《中国数学简

史》(*Chinese Mathematics – A Concise History*, 由 John N. Crossley 和 Anthony W. C. Lun 译为英文, Oxford: Clarendon press, 1987 年). 另一较早的著作是三上义夫(Yoshio Mikami) 的 *The Development of Mathematics in China and Japan*, New York: Chelsea, 1974. 李倍始(Ulrich Libbrecht) 所著的 *Chinese Mathematics in the Thirteenth Century: The Shu – shu Chiu – chang of Ch'in Chiu – shao*, (Cambridge: MIT Press, 1973) 中对 13 世纪中国数学的各个方面及其与其它时期数学的关系以及在其它国家的传播与影响作了详细的讨论. Frank Swetz 和 Ang Tian Se 的 A Brief Chronological and Bibliographic Guide to the History of Chinese Mathematics(*Historia Mathematica*, 11, 1984, 39 – 56) 一文中给出了中国数学史的文献指南. 有关印度数学通史的著作有 B. Datta 和 A. N. Singh 的 *History of Hindu Mathematics*(Bombay: Asia Publishing House, 1961) 和 C. N. Srinivasiengar 的 *The History of Ancient Indian Mathematics*(Calcutta: The World Press Private Ltd., 1967). 在 Walter E. Clark 所译的 *The Āryabatīya of Āryabhata*(Chicago: University of Chicago Press, 1930) 一书中可以找到阿耶波多著作的一些原文, 在 *Algebra with Arithmetic and Mensuration from the Sanskrit of Brahmagupta and Bhāskara*(London: John Murray, 1817) 一书中可以找到由 H. T. Colebrooks 所译的波罗摩笈多和巴什迦罗的主要数学文献.

1. E. l. Konantz, "The precious mirror of the four elements," *China Journal of Arts and Science* 2(1924), 304 – 310. 引自 Chien Chiu Shimoju 的朱世杰工作介绍.
2. Frank J. Swetz, *The Sealand Mathematical Manual: Surveying and Mathematics in Ancient China*(University Park, Pa.: Pennsylvania State University Press, 1992), p. 20. 这本小册子包括了《海岛算经》完整的译文和分析, 还有对古代中国测量学的讨论.
3. 同上.
4. 有关中国三角学工作的更多内容参见 Yabuuti Kiyosi, "Researches on the *Chiu – chih li* – Indian Astronomy under the T'ang Dynasty," *Acta Asiatica* 36 (1979), 7 – 48, 以及 Christopher Cullen, "An English Century Chinese Table of Tangents," *Chinese Science* 5(1982), 1 – 33. 前一篇文章包括对《九执历》的英文翻译和评论, 而第二篇文章则是详细给出了僧一行推导其正切表的方法和动机.
5. Libbrecht, *Chinese Mathematics*, p. 269.
6. 同上, p. 227.
7. Libbrecht, *Chinese Mathematics*, p. 62.
8. 同上, 31. 此外 Libbrecht 引用了秦的同时代人周密的说法.
9. 最早给出秦九韶方法成立的证明是在 V. A. Lebèsgue 的 *Exercises d'analyse numérique*(Paris, 1895), 第 56 页中. Kurt Mahler 在"On the Chinese Remainder Theorem"(*Mathematiche Nachrichten* 18(1958)) 一文中给出了现代的证明.
10. 李俨和杜石然, *Chinese Mathematics*, p. 163.
11. 详细内容和事例参见 Libbrecht, *Chinese Mathematics*, 第 17 章.
12. 蓝丽蓉(Lam Lay Yong), "On the existing Frangments of Yang Hui's Hsiang Chieh Suan Fa," *Archive for History of Exact Sciencs* 5(1969), 82 – 86. 有关应用巴斯加三角形解方程的详细内容参见 Lam Lay Yong, "The Chinese Connection between the Pascal Triangle and the Solution of Numerical Equations of Any Degree," *Historia Mathematica* 7(1980), 407 – 424.
13. Lam Lay Yong, "Li Yie and his Yi Gu Yan Duan," *Archive for History of Exact Sciencs* 29(1984), 237 – 266.
14. 问题引自: Jock Hoe, *Les systemes d'equations polynomes dans le Siyuan Yujian*(1303)(Paris: Collège de France, Institut des Hautes études Chinoises, 1977). pp. 94ff.
15. 阿耶波多的命题译自 W. E. Clark, *The Āryabatīya*, 芝加哥大学 1930 年版权, 经允许后重印.
16. E. Burgess, "Tanslations of the Sūrya – Siddhānta, a Textbook of Hindu Astronomy," *Journal of the American Oriental Society* 6(1860), 141 – 498, p. 252.
17. 引自: R. C. Gupta, "Bhāskara I's Appooximation to Sine," *Indian Journal of History of Science* 2(1967), 121 – 136, p. 122.
18. 在 R. C. Gupta, "On Derivation of Bhāskara I's Formula for Sine," *Ganita Bhāratī* 8(1986), 39 – 41 中讨论了这一提法, 在 Kim Plofker 的博士论文 *Mathematical Approximation by Transformation of Sine Functions in Medieval Sanskrit Astronomical Text*(Brown University, 1995) 中更详细的讨论了这种估值方法.

19. H.T. Colebrooks, *Algebra with Arithmetic*, pp. 325ff.

20. 参见 B.L. Van der Waerdwn, *Geometry and Algebra in Ancient Civilization*, *chapter* 5A.

21. H.T. Colebrooks, *Algebra with Arithmetic*, pp. 364ff.

22. 同上, pp. 175ff.

23. 参见 C.N. Srinavasiengar, *History of Ancient* 第 113 页的证明.

24. Krishnaswami A.A. Ayyangar, *Journal of the Indian Mathematics Society* 18(1929), 232 - 245.

25. C.O. Selenius, "Rationale of the Chakravāla Process of Jayadeva and Bhāskara Ⅱ," *Historia Mathematica* 2(1975), 167 - 184 中对解佩尔方程的完整过程作了详细的讨论.

26. 参见 Plofker, *Mathematical Approximation*, 由于版权原因, 这里不能给出该问题的更多细节.

27. H.T. Colebrooks, *Algebra with Arithmetic*, pp. 339 - 340.

28. 同上, p. 340.

29. 同上, p. 135.

30. 同上, p. 19.

31. 同上, p. 40.

32. 同上, p. 7.

33. 同上, pp. 346 - 347.

34. 同上, pp. 207 - 208.

35. 同上, pp. 215 - 216.

36. Mahāvīra, *Ganitasārasangraha*, M. Rangācārya, ed. and trans. (Madras: Government Press, 1912), p. 134.

37. H.T. Colebrooks, *Algebra with Arithmetic*, p. 235.

38. Gurugovinda Chakravarti, "Growth and Development of Permutations and Combinations in India," *Bulletin of Calcutta Mathematical Society* 24(1932), 79 - 88. 更多的内容参见 N.L. Biggs, "The Roots of Combinatorics," *Historia Mathematica* 6(1979), 109 - 163.

39. 引自: *Brhat Samhitā*, *Chapter* 77, rule 20, 英译文见: J.H. Kern, "The Brhatsamhita of Varahamihira," *Journal of Royal Asiatic Society* (1975), 81 - 134.

40. Mahāvīra, *Ganitasārasangraha*, sec. 1.

41. 参见: R.C. Gupta, "Varāhamihira's Calculation of ^nC_r and the Discovery of Pascal's Triangle," *Ganita - Bhāratī* 14(1992), 45 - 49.

42. Mahāvīra, *Ganitasārasangraha*, p. 150.

43. Colebrooks, *Algebra with Arithmetic*, p. 49.

44. 同上, p. 124.

45. Clark, *The āryabhatīya*, p. 21.

46. 引自: D.E. Smith, *History of Mathematics* (New York: Dover, 1958), Vol. 1, p. 166. 关于印度数学符号的更多内容详见: Saradakanta Ganguli, "The Indian Origin of the Modern Place - Value Arithmetical Notation," *American Mathematical Monthly* 39(1932), 251 - 256, 389 - 393, 和 40(1933), 25 - 31, 154 - 157.

47. Yabuuti Kiyosi, "Researches on the *Chiu - chih Li*," p. 12.

48. 参见 Lam Lay Yong, "The Conceptual Origins of our Numeral System and Symbolic Form of Algebra," *Archive for History of Exact Sciencs* 36 (1986), 184 - 195. 以及 "A Chinese Genesis: Rewriting the History of Our Numeral System," *Archive for History of Exact Sciencs* 38(1988), 101 - 108. 更多的详细内容见 Lam Lay Yong 和 Ang Tian Se 的著作 *Fleeting Footsteps: Tracing the Conception of Arithmetic and Algebra in Ancient China* (Singapore: World Scientific, 1992).

49. Frank J. Swetz, *The Sealand Mathematical Manual*, p. 21.

50. Frank J. Swetz and T.I. Kao, *Was Pythagoras Chinese?*, p. 60.

51. Mahāvīra, *Ganitasārasangraha*, p. 73.

中世纪中国和印度数学概览

3世纪	刘徽(Liu Hui)	测量数学
3世纪后期	孙子(Sun Zi)	中国剩余定理
5世纪	张邱建(Zhang Quijian)	百鸡问题
5世纪后期	阿耶波多(Āryabhata)	正弦表、数学方法
6世纪	彘日(Varāhamihira)	组合规则
7世纪	波罗摩笈多(Brahmagupta)	不定方程
8世纪早期	瞿坛悉达(Chutan Hsita)	中国的印度正弦表
683—727	一行(Yi Xing)	正切表
9世纪	马哈维拉(Mahāvīra)	代数及组合问题
11世纪	贾宪(Jia Xian)	巴斯加三角形
1114—1185	巴什迦罗(Bhāskara)	佩尔方程
1192—1279	李冶(Li Ye)	几何问题的代数方程
1202—1261	秦九韶(Qin Jiushao)	线性同余、解方程的技巧
13世纪后期	杨辉(Yang Hui)	解方程的技巧
13世纪后期	朱世杰(Zhu Shijie)	方程组
1552—1610	利玛窦(Matteo Ricci)	将西方数学传入中国

第7章 伊斯兰数学

> 你们清楚知道…… 我为什么开始寻求证明由古希腊人提出的许多论断的理由…… 还有我对于这门学科所感受到的那种激情…… 你们为此责备我专注于这些几何方面的章节而不了解这些课题的真正要害恰恰在于在每个地方它们都逾越了其应有的范围…… 无论他[几何学家]采取什么方式，通过尘世的活动直至汲取神的教喻，都不能达到目标，这是因为难于理解它们的含意…… 也是因为并非每个人都能对它们的概念有所了解，特别对那些把证明技巧拒之门外的人更是如此.
>
> —— 公元1030年比鲁尼《求圆的弦长》一书的序言[1]

有一个故事说，奥玛·海亚姆在做学生时与他的两个同学尼扎姆·穆勒克和哈桑·萨巴赫曾有个协议，其大意是说，其中哪一个先取得高官或发了大财就要帮助另外两个人. 实际上是尼扎姆成了塞尔柱苏丹杰拉勒丁·马利克沙的重臣. 他实践了诺言. 哈桑得到了官廷司库的职位，但在他试图取代他的朋友所受苏丹宠信的地位后便被逐出了官廷. 另一方面，奥玛尔拒绝了高官职位，转而接受了一个有适当薪俸的位置，让他有闲暇从事研究和写作.

在7世纪的前半叶，出现了一个新的阿拉伯文明. 在先知穆罕默德的感召下，一种新的神教宗教伊斯兰教发展起来，并很快得到了阿拉伯半岛上居民的忠诚和拥戴. 在630年穆罕默德占领麦加后不到一个世纪的时间里，随着新宗教先是在中东信仰多神的部落中的传播，而后又在其他信仰的信徒中的传播，伊斯兰军队征服了一大片的领土. 首先把叙利亚，然后把埃及从拜占庭帝国中夺取过来. 波斯则在642年被征服，不久，战无不胜的军队到达了远至印度及中亚的部分地区. 在西面，北非很快就被侵占了，在711年伊斯兰军队又进入了西班牙. 他们前进的步伐最终于732年被查理·易特的军队在图尔阻止. 于是征服的问题便让位于新的问题，即统治和管理庞大的新帝国. 穆罕默德的继承者们即那些哈里发们，最先把首都建在大马士革，但是在百多年的战争，既包括了伟大的胜利也包括了惨重的失败之后，哈里发的辖地分裂成了许多部分. 在阿拔斯哈里发统治下的东部地区，奢华生活的增长及征服战争的中止创造了新文化发展的有利条件.

在 766 年，曼苏尔哈里发把他的首都设在了巴格达，它很快便成了一个繁荣的商业中心和智力中心. 正宗伊斯兰教的初期冲动很快被更具宽容的氛围所替代，并且辖区内全体居民的智力成就都受到了欢迎. 从 786 年到 809 年的统治者哈伦 · 赖世德哈里发在巴格达建立了一座图书馆. 从近东地区各类学术机构收集了大量手稿，这些机构是由那些为逃避古代雅典和亚历山大学术界迫害的学者们建立的. 这些手稿中包括了许多古希腊的数学和科学的文章. 一个把它们翻译成阿拉伯文的计划很快就开始了. 哈伦的继承人马蒙阿里发(813—833) 创建了一个研究所"智慧宫"，它一直存在了二百多年. 在哈里发辖区的各个地区的学者受邀到这个研究所进行翻译，把希腊的和印度的著作翻成阿拉伯文，同时还进行了最初的研究工作. 到了 9 世纪末，欧几里得，阿基米德，阿波罗尼奥斯，丢番图，托勒密及其他希腊数学家的许多主要著作都已翻译成了阿拉伯文，并可供给集中在巴格达的学者们作研学之用. 伊斯兰学者们还吸取了当时仍存在于底格里斯 – 幼发拉底河谷的古代数学的巴比伦传统书写法，另外还学习了印度人的三角学.

伊斯兰学者所做的远不只是简单地把这些来源汇拢一起. 他们把它们混合成一个新的整体，特别地像本节开始所引用的文字中所表明的那样，他们使他们的数学中浸满了他们所信奉的神的灵感. 过去年代中，那些有创造性的数学家总是使研究大大超越了当前的需要，但是在伊斯兰中许多人感到这只是真主的要求. 至少是在其初期，伊斯兰文化并不把"世俗知识" 视为与"神赐文化" 相冲突，而是当作通向后者的一条道路. 因而学术得到鼓励，那些被证明具创造火花的人们常常得到统治者(通常是世俗和宗教当局双方) 的支持，从而他们能够尽可能地追循他们自己的想法. 数学家的回报则是在他们著作的开端和结尾总要祈求神的保佑，甚至整个正文中有时也提到神的帮助. 另外，由于统治者对日常生活需要的兴趣，伊斯兰数学家们不仅要在理论上作出成绩，而且还要对实际应用有所贡献.[2](当然，也有一些伊斯兰统治者并不支持数学和科学的研究，他们相信需要知道的所有东西都已经写在可兰经中了.)

知道了伊斯兰在总体上对科学，特别是数学的影响后，这个时期的数学在这里被指称为"伊斯兰的" 而不是"阿拉伯的" 便是有根据的了，尽管并非那时所有的数学家都是穆斯林. 但是，在伊斯兰区域中通常都使用阿拉伯语言，从而我们要讨论的那些著作也都是用这种语言写的. 现在还不能写成一部完整的中世纪伊斯兰的数学史，这是因为还有如此多的这些阿拉伯文手稿没有被研究过甚至没有被阅读过，它们仍然躺在全世界的图书馆里. 由于越来越多的材料被编译，近来情况有所改善，但是政治上的困难使得许多重要的收集工作继续被阻挡在外. 然而，伊斯兰数学的总体轮廓已经清楚了. 特别地了解到伊斯兰的数学家已充分把十进位值制数系发展到十进制分数，把代数的研究系统化，并开始考虑代数与几何之间的关系，把从印度取来的组合数学的规则重新作成了一个抽象的体系，研究并推进了大部分的欧几里得、阿基米德和阿波罗里奥斯的希腊几何论著，并且在平面和球面三角学上作出了重大的改进(图 7.1).



图 7.1　突尼斯邮票中的阿拉伯的科学贡献.

7.1　十进制算术

十进位值制数系在 7 世纪中期由印度至少传播到了叙利亚. 可以肯定的事实是，在智慧宫创建的时候它已经存在于伊斯兰的领地了. 事实上在 773 年，一位印度学者造访了巴格达的曼苏尔宫廷，他随身带来了一本印度的天文学教科书，很可能是婆罗摩笈多的《婆罗摩修正体系》. 这位哈里

发下令将它翻译成阿拉伯文.此书除包含了印度的天文系统外,至少还提及了印度的数字系统.但是穆斯林们已经有了一个满足于数学应用的数字系统.事实上,有两个系统正在那里使用.市场的商人们一般使用手指计算的形式,它是祖祖辈辈留传下来的.在这个系统中,计算通常是以心算形式进行.数字由文字来表达,而分数则由六十进制的巴比伦度量来表示.当数必须写下来时,便使用了一种字母系统,用阿拉伯字母来代表数字.在8世纪到13世纪间写成了许多讨论这种或那种数系的阿拉伯算术教科书.

印度数系的知识逐渐地渗透进了伊斯兰数学.最早出现的处理印度数系的算术教科书是《依照印度人方法做加法和减法的书》,它是由默罕默德·花拉子米(Muhammad ibn-Mūsā al-Khwārizmī)(780—850)所著,他是智慧宫的一位早期成员(见补遗 7.1).可惜,已没有这本著作的阿拉伯文手稿存世了,而只有12世纪中在欧洲出版的许多不同的拉丁文版本.在这本书中花拉子米介绍了代表了前九个数的九个字符,以及如拉丁文版告诉我们的那样,以一个圆圈代表了零.他指出了如何用这些字符把任意一个数写成我们所熟悉的那种位值记号.然后他描述了各种算法:加、减、乘、除、一半、两倍及确定平方根,并给出使用的例子.但是这些算法通常都是在一个铺上沙子的土板上书写的.因此计算都通常设想成在进行下一步运算时便把上一次运算写下的数字从沙盘上抹去,直至得到最后的答案.虽说花拉子米并非是一贯地,却是最经常地按埃及人的样子把分式表示为单位分数的和.一个值得注意的重要之处是我们的位值系中最重要特征即十进位小数仍然没有出现.尽管如此,花拉子米的工作不但在伊斯兰世界中是主要的,而且也因为它对许多欧洲人介绍了十进位值系统而具有重要性(补遗 7.2).

跨越了整个下世纪的大量阿拉伯文的算术著作解释了印度方法,这些著作既写到了他们自己的东西也联系到了前面所提到过的那些更古老的系统.现存最早的阿拉伯算术书是乌格里狄西(Abu l-Hasan al-Uqlīdīsī)所写的《印度算术》,写于952年的大马士革.作者清楚阐明了他所认识到的印度数系会最终取得成功的一个主要理由:

> 大多数文书不得不使用它(指印度记数法),因为它容易,快捷,不需多少准备,不要多少时间便能得到答案,也不需在心里忙于计算,即必须盯住自己的双手使得在他讲话时不至于破坏他的计算;并且当他离开去做别的事再转回来时,他将发现它仍是原样,因而可以继续算下去而免去了记忆的麻烦,从而又继续做他的事了.其它的(算术)就不一样了,需要扳弄手指并做其他必须做的事.大多数的计算者不得不使用它(印度方法)来处理那些不能用手处理的数,因为它们实在太大了.[3]

补遗 7.1 阿拉伯姓名

在我们最初提及一个伊斯兰数学家时,我们给出了他的全名,而在之后为了方便则给出了省略的名字.注意,阿拉伯的名字所包含的不仅有这个人的教名,而且也会包含他与直系家族中一个或多个人员的关系("ibn"表示"…… 的儿子"),他或他的先人的出生地,他儿子的名字("abū"表示"…… 的父亲"),以及一个或多个表示某种特性的称呼.例如,乌格里狄西(al-Uqlīdīsī)表示与欧几里得有关.换句话说,有这个名称的数学家可能是欧几里得著作的阿拉伯文的抄印者.

补遗 7.2	由阿拉伯文来的数学词汇
	有三个英文数学字眼大概来自花拉子米的算术书. 这本书的一个拉丁文文稿以"花拉子米说" 开始, 即"Dixit Algorismi". 经过一些误解和差讹, "algorismi" 不久便成了表示各种算术运算的一个词汇且最终成了英文字"algorithm". 我们的"zero" 大概是由阿拉伯文 sifr 来的, 它被拉丁化为"zephirum". sifr 是梵文 sūnyā 的阿拉伯译文, 其意思是"空". sifr 的另一个中古希腊语的翻译是"cifra", 它变成了现代英文"cipher".

像花拉子米一样,乌格里狄西的书也处理了各种算术的算法. 但它有两项重要的革新. 首先作者指出了如何在纸上进行算术计算. 正如他所注意到的那样,一些人以为"看到手持土板的书记员 …… 坐在市场里是令人厌恶的 …… 故而我们要以某种东西来替代它." 例如,乌格里狄西给出了 3249 乘以 2735 的下列计算过程. 他把第一个数字写在第二个数的上面,然后把第一个的每个数码乘以整个第二个数,再把所得到的这些项加起来. 例如该计算的第一行是 6 21 9 15(= 2 · 3,7 · 3,3 · 3,5 · 3),

			3249			
			2735			
6	21	9	15			
	4	14	6	10		
		8	28	12	20	
			18	63	27	45

所要答案 8 886 015 由沿着各个不同位置小心地逐列相加得到. 于是答数中右数第二个数字由 20 和 27 中的 0 与 7 加上 45 中的 4 得到. 右数第三个数字由前次计算"进位" 的 1 加上 20 中的 2,27 中的 2,10 中的 0,12 中的 2 和 63 中的 3 都可以验算它们.

其次,乌格里狄西处理了小数,这些小数是在中国之外最早有纪录的例子. 对它的处理出现在乌格里狄西的关于取半的章节中:"在数的原理的设计中,任意位置上的 1 的一半是在它前面得 5. 因此,如果我们把一个奇数取半的话,我们则在它前面把这一半置为 5,个位用在它上标以符号 ′ 表示. 而个位则变成它以前的十位位置. 下一步,我们将 5 一半就如同按惯例取成了百位. 这样一直做下去."[4] 这里的小数的中心思想是清晰的. 在处理小于 1 的数时,人们完全依对整数的运算一样来进行. 只是在完成运算后人们还需烦心一下小数的位置. 乌格里狄西给出了对 19 连续取 5 次一半的例子. 依次地,他得到 9′5,4′75,2′375,1′1875 和 0′59375. 他把这最后的数读作十万分之 59375. 类似地,在关于增数的章节中,他发现求一个数的十分之一只需简单地把它"往下移一位" 重写即可. 因此把 135 连续增加自身十分之一五次时,他写出了

1	3	5	
	1	3	5

其和是 148′5. 此数的十分之一是 14′85; 新的和则是 163′35. 继续此过程三次最后的答案是 217′41885.

虽然乌格里狄西使用了小数,但并不清楚他是否完全掌握了它们的意义. 他处理的只是除以 2 和 10 的除法,而他并没有计算过例如 $\frac{14}{3}$ 的小数形式. 相比较而言,塞毛艾勒(al-Samaw′al ibn Yahyā ibn Yahūda al-Maghribī, 1125—1174),在他 1172 年的书《论算术》中表明他充分了解了小数,这是他在写有关近似的内容时涉及到的. 他从描述基本思想开始:"给定比例的位置先从个位开始,随后按

十分之一的比例相继地一直下去,因此我们假定按同样的比例在个位的另一边上相继跟随着十的部分,而且个位位于两个部分之间,其中一部分是整数的位置,它的位数以同样的方式无限移动下去,而另一部分是无限可除的."[5]

作为例子,塞毛艾勒举出了210除以13并注意到这个除法不会除尽,因此,可以进行到任意远的地方.他写到五位的结果是16加上10中的1份加上100中的5份加上1000中的3份加上10 000中的8份再加100 000中的4份.相似地,他计算了10的平方根为3加上10中的一份加100中的6份加1000中的2份加10 000中的2份加100 000中的7份加1 000 000中的7份(3.162 277).与他的先辈不同,他仍旧用文字来描述各个位置.但是他懂得用小数来表示有理数或无理数的值.事实上,塞毛艾勒用了一种类似于在中国使用的方法来计算高次根,他清楚地注意到这个算法的逐步计算的目的:"因此我们作运算以确定立方体的边,正方 - 正方,正方 - 立方及其他[幂次]方的边.这个方法使我们……得到无限多个答案,每一个都比前面一个离真实答案更精确,更接近."[6] 塞毛艾勒至少在理论上意识到人们可以计算出一个数的无限小数展开式,而此展开式的有限小数"收敛"于这个确定值,而这个确定值是不能用有限形式来表达的.

但是,即便是这个重要著作,也不能把位值系统完全展开.正是在15世纪初卡西(Ghiyāth al-Dīn Jamshīd al-Kāshī)在1429年的著作中,我们第一次看到了同时完全掌握了小数的思想和对它们的便捷的记号(图7.2).卡西用一条竖线来隔开一个数的整数部分和它的小数部分.那么这时我们可以说印度 - 阿拉伯位值系统已经完成了.这个系统,包括小数在内,也在这个时间左右出现在拜占廷的一本教科书中,书中把这个方法描述为"土耳其的"也就是说是伊斯兰的.这本教科书是在1562年带到了威尼斯,但甚至在此之前,这种记号有时也已出现在欧洲的著作中了.然而直到17世纪初,这个完整的系统才在欧洲被使用.



图7.2　伊朗邮票上的卡西.

7.2 代　数

伊斯兰数学家最重要的贡献在代数领域.他们把巴比伦人已发展了的材料结合经典的希腊几何遗产产生出一个新的代数学,并对它进行了推广扩张.在9世纪末之前,主要的希腊数学经典著作在伊斯兰世界已广为人知.伊斯兰学者们对它们进行了研究并写了关于它们的评注.他们从学习这些希腊著作中得到的最重要的思想是关于证明的观念.他们吸收了这样的思想,即除非人们能够证明一个数学问题的解是成立的,否则便不能认为这个问题已经被解答.人们应该如何证明这一点,特别是对一个代数问题?答案似乎是清楚的.惟一真实的证明是几何的.毕竟在希腊书中找到的是几何而非代数.因此伊斯兰学者们赋于自己的任务是证明代数规则的合理性,这些代数规则或是古老的巴比伦人的或者是他们自己发现的,而且是通过几何来证实它们的合理性.

7.2.1 花拉子米和伊本·吐克的代数

最早的伊斯兰代数教程之一是由花拉子米在825年写成的,书名是《al-Jabr与al-Muqabala计算概要》,它最终的影响力比起他的算术著作还大.al-Jabr这个词可以翻译为"还原",它所指的意思是把方程式一边的被减量"转移"到另一边,把它变成一个被加量.al-muqabala可翻译为"比较",是指把一个正项通过同时在方程式两边减去相等的量所进行的简约.因此从$3x+2=4-2x$到$5x+2=4$的转化是个al-Jabr的例子,而将后式转化为$5x=2$则是al-muqabala的例子.我们的英文代数这

个字 algebra 是阿拉伯文 al-Jabr 的讹用，当这本书和其他类似的专著被翻译为拉丁文时就开始使用这个词了.并没有把 al-Jabr 这个词翻译过来，却把它当作了这门学科的名字了.

在他的序言中花拉子米解释了他为什么想到要写这本书：

> 对科学的喜爱，而真主正是根据科学选择了忠诚教徒们的统领依玛目马蒙 …… 他对于有学问的人表现出的谦逊和屈尊，他保护和支持他们在克服困难方面所显示的果断和迅捷，所有这些鼓励了我去创作一部短小的著作以用 al-jabr 和 al-muqabala 去作计算，并把它限于那些算术中最容易和最有用的部分，诸如人们在继承遗产，分割财物，法律诉讼，贸易以及他们的其它一切交往中所经常需要的方面，或者是在需要丈量土地，挖掘运河，几何计算以及其他各式各样的方面.[7]

花拉子米的兴趣在于写一本实用手册而不是一本理论书籍.但是在智慧宫所引进的希腊数学已充分地影响到他，使得在这样一本手册里他也感到不得不对他的代数进程给出几何的证明.然而这种证明并非是希腊式的.事实上，它们看起来非常类似于巴比伦人的几何推断，而正是由于这种推断产生了代数的算法.像他的东方先行者一样，花拉子米也给出了大量的例子和问题，但希腊的影响通过他对想要解决的问题进行系统分类显现了出来，也在对他的方法的非常详尽的解释中显现出来.

一开始花拉子米便注意到"人们一般在计算中所需要的是 …… 一个数，"[8] 即一个方程式的解.因此这本书便成为一本解方程的手册.他所处理的量一般为三种，即（未知量的）平方，这个平方的平方根（即未知量本身）以及纯粹的数（即方程式中的常数）.然后他留意到，用这三种量可以写出六种类型的方程式：

1. 平方等于根（$ax^2 = bx$）.
2. 平方等于数（$ax^2 = c$）.
3. 根等于数（$bx = c$）.
4. 平方和根等于数（$ax^2 + bx = c$）.
5. 平方和数等于根（$ax^2 + c = bx$）.
6. 根和数等于平方（$bx + c = ax^2$）.

人物小传

花拉子米(780—850)(Muhammad ibn mūsā Al-Khwārizmī)

花拉子米（图7.3）或者是他的祖辈，应是来自花拉子模，这是咸海南部的一个地方，现在是乌兹别克斯坦和土库曼斯坦的辖地.花拉子米是马蒙哈里发创建的智慧宫的第一批学者之一，而且是在公元847年应召为瓦舍格哈里发进行临终星象占卜的天文学家之一.传说花拉子米曾保证哈里发会再活五十年，然而事实是在十天之后他就死了.或许这是由于花拉子米感到做一个把坏消息带给他的统治者的人是件坏事的缘故才这样说的吧.除了本文中详细叙述的数学方面的贡献外，花拉子米还写了一本地理方面的书，在其中他展示了一幅伊斯兰世界的地图，它远远优于从托勒密的著作中所知道的.



图7.3 前苏联邮票上的花拉子米.

之所以作出这种六层分类是因为伊斯兰数学家不像印度人，他们不处理负数.在他们的系统中，方程式的系数以及根都同样必须是正的.所列出的这些类型只是具有正解的方程式.我们的标

准形式 $ax^2+bx+c=0$ 对于花拉子米说来毫无意义,因为如果所有的系数都是正的,根就不可能是正的了.

对前三类方程式花拉子米的解法是直接了当的.我们只需注意到零并不被当作第一类方程式的解.对于复合型方程式,他的规则更有意思.我们来展示他对类型 4 的解法.因为花拉子米不使用符号,那么我们将照他那样用文字来写出所有的一切,包括他例子中的数:“平方以及它的根的十倍加起来为 39 会是什么样呢?答案是:你将根的倍数取半,现在的例子中便是 5.你将这数自乘,其积为 25.把它加到 39 上,和为 64.现在取其根得 8,然后从它中减去根的倍数的一半即 5,最后余下 3.这就是你要找的平方的根.”[9]

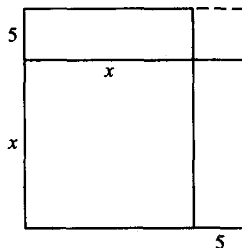


图 7.4 花拉子米所给出的对 $x^2+10x=39$ 的解正确性的几何说明.

花拉子米对他的步骤的文字描述在本质上与巴比伦的书记员所作描述是一样的.用现代的记号表示即 $x^2+bx=c$ 的解是

$$x=\sqrt{\left(\frac{b}{2}\right)^2+c}-\frac{b}{2}.$$

花拉子米对此步骤的几何式的合理性证明也展示出了他的巴比伦传统.先作一个代表 x^2 的正方形,然后添加在它上面两个矩形,它们的宽都为 5(“根的倍数的一半”)(图 7.4).正方形面积和这两个矩形面积之和于是为 $x^2+10x=39$.人们现在可以补上一个面积为 25 的正方形使得整个图形成为一个正方形,其总面积为 64.于是便容易解出 $x=3$.这个几何描述对应于巴比伦人对 $x^2+\frac{4}{3}x=\frac{11}{12}$ 的解的描述(见第 1 章及图 1.29).

虽然花拉子米的几何描述方法看起来是取自于巴比伦,他或他的在这个领域中的(无名的)先辈们成功地把解二次方程的中心从实际上求正方形的边长改变为求满足某些条件的数.例如,他并不把“根”这个词解释为正方形的边而是当作“任何一个由单元组合成的可以自己乘自己的事物,或者是乘上自己大于 1 的任何数或自乘后变得小于 1 的数.”[10] 当平方项的系数不是 1 的时候,他解类型 4 的二次方程式的步骤是,首先用算术方法乘以或除以适当的数使原来的系数变为 1,然后再像前面一样进行.在后来的书中,当在讨论“多项式” $100+x^2-20x$ 与 $50+10x-2x^2$ 的相加时,花拉子米甚至承认:“这是不容许有任何图形的,因为存在三个不同的类,即平方、根和数,没有什么图形能对应并表示它们……(但是)用文字来阐述则是容易的.”[11]

最后,对类型 5 的方程式即平方和数等于根的解法和几何描述,表明他能够处理具有两个正根的方程式,至少对数值方程能够做到,这不同于巴比伦人.这时方程式为 $x^2+c=bx$,他对解的步骤的文字描述很容易翻译成我们的公式

$$x=\frac{b}{2}\pm\sqrt{\left(\frac{b}{2}\right)^2-c}. \tag{7.1}$$

事实上,他说人们可以应用加法或减法来得到一个根,而且注意到了有解的条件:“如果[根的倍数的一半自乘的]积小于与平方相关联的那个数,则这个例题是不可能有解的,但是如果此积等于那个数,则平方的根便只等于根的倍数的一半,并不需要用加法或减法.”[12] 对这种情形的几何证明使我们想起了巴比伦人对方程组 $x+y=b$, $xy=c$ 的描述,他们会将此方程转换成这个方程组(见第一章,图 1.27),这时只需在等式(7.1)中处理减法就行了.在图 7.5 中,正方形 $ABCD$ 代表 x^2,矩形 $ABNH$ 代表 c.因此 HC 代表了 b. G 平分 HC,延长 TG 到 K 使 $GK=GA$,并作成矩形 $GKMH$.现在可以清楚看出矩形 $MLRH$ 等于矩形 $GATB$.由于正方形 $KMNT$ 的面积为 $\left(\frac{b}{2}\right)^2$,此正方形去掉正方形

$KLRG$ 等于矩形 $ABNH$ 或 c，从而得到正方形 $KLRG$ 的面积为 $\left(\frac{b}{2}\right)^2 - c$，因为此正方形的边等于 AG，从而由等式(7.1)利用减号给出的 $x = AC = CG - AG$ 便得到了答案. 虽然花拉子米简短地提到 CR 也可以表示一个解，但他并没有以一个图示来证明它，也没有在他的图示中处理他在文字叙述中提到的特殊条件.

花拉子米这部著作的标题中包含有"概要"这个词，它提示我们存在有其他一些书籍，它们给出了代数步骤的更详细讨论及其相伴的几何证明. 然而尚存的只有这样一本著作的片段，即一本较长的著作《代数学》(Kitāb al-jabr wa'l muqābala) 中的一节"混合型方程式中的逻辑要点"，这书是伊本·吐克(Abd al-Hamīd ibn Wāsi ibn Turk al-Jīlī) 所著，他是花拉子米的同时代人，对他的生平几乎一无所知. 甚至对伊本·吐克是来自伊朗还是阿富汗还是叙利亚也各有说法.

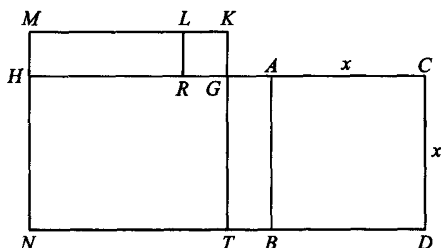


图 7.5　花拉子米对 $x^2 + c = bx$ 解的几何证明.

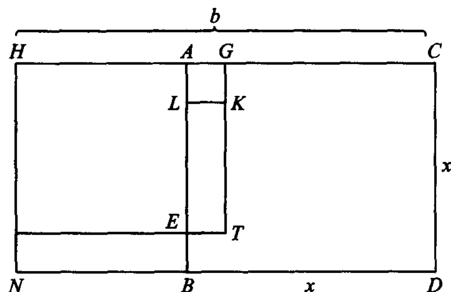


图 7.6　伊本·吐克对 $x^2 + c = bx$ 的一种情形的几何证明.

无论如何，伊本·吐克著作的残存章节处理了花拉子米的1,4,5和6型的二次方程式，并且包含了远较花拉子米著作中更为详细的对解法的几何描述. 特别在类型5的情形，伊本·吐克对所有的可能情形都给出了几何解释. 他的第一个例子与花拉子米的相同，即 $x^2 + 21 = 10x$，但是他注意到 CH 的中点 G 既可以像花拉子米的图示那样在线段 AH 上，也可以在图7.6中的线段 CA 上，并依此来进行几何证明. 这时，类似于图7.5那里的样子，其中正方形和矩形都是完整的，但是解 $x = AC$ 由 $CG + GA$ 给出，从而用了等式(7.1)中的加号. 另外，伊本·吐克讨论了他所谓的"中间情形"，这时平方的根正好等于根的倍数的一半. 他对此情形的例子是 $x^2 + 25 = 10x$；其几何图示则简单地由分成两个相等正方形的矩形组成.

伊本·吐克进一步注意到"当数值……大于根的倍数的一半(的平方)时，这类方程之不可解有其逻辑的必然性，"[13] 例如，像在 $x^2 + 30 = 10x$ 的情形. 他又一次诉诸于几何推断. 假定 G 位于线段 AH 上，像前面那样我们知道正方形 $KMNT$ 大于矩形 $HABN$(图7.7). 但是问题的条件却表明后一个矩形等于30而前一个只等于25. 对 G 位于 CA 上的情形相似的推断也成立.

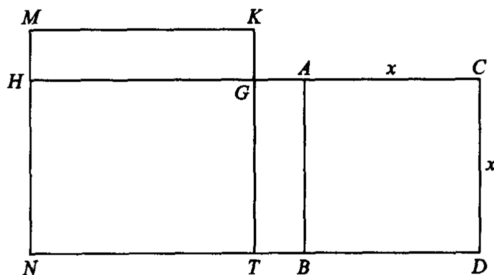


图 7.7　对 $x^2 + 30 = 10x$ 的不可能解的伊本·吐克的几何证明.

虽然关于二次方程这一节是伊本·吐克的代数中惟一尚存的，但花拉子米的书却包含了许多其他的有意思的内容，包括参照处理数的类似方式引进了对代数表达式的处理. 例如，他知道了如果 $a \pm b$ 乘以 $c \pm d$ 则必有四个积. 虽然他的数没有一个是

负的，但他的确了解处理乘法和符号的规律. 正如他所说，“如果各单元(在我们的记号中即 b 与 d)…… 为正的，则最后一个积为正；如果它们都为负则第四个积同样为正. 但是如果它们中一个为正而另一个为负，则第四个积为负.”[14]

花拉子米的书还收集了大量的问题. 其中许多问题涉及到这样的处理方法，而其中大多数问题则导致了解二次方程式. 例如，其中一个问题说，“我把 10 分成了两部分，并将每个部分自乘，再将它们放在一起，并对它们加上未作乘积前的那两部分的差，所有这些总共等于 54”[15]. 不难把此问题翻译为方程式 $(10-x)^2+x^2+(10-x)-x=54$. 作者将此简约为方程 $x^2+28=11x$，然后用他的对第五类方程的规则得到 $x=4$. 但忽略了第二个根 $x=7$，因为这时这两个平方的和会得出 58，从而问题的条件不能满足. 在另一个例子中，花拉子米处理了非有理根：“我将 10 分成两部分；将第一部分乘以 10 而将第二部分自乘，且这两个积相等.”[16] 这里的方程式是 $10x=(10-x)^2$，解为 $x=15-\sqrt{125}$. 他在这里仍然略去了带正号的那个解，因为 $15+\sqrt{125}$ 不可能是 10 的一“部分”.

虽然花拉子米在他的序言中许诺他将写那些“有用”的东西，但他那些导致二次方程的问题却很少与“实用”有关. 它们中的许多问题都类似于前面的那些例子，总以“我把 10 分成两部分”作为开始. 有几个问题与在一群人中分钱有关，但即便是这些问题也没有什么实用意义. 事实上，其中一个问题归结到方程式 $x^2+x=3/4$，其中 x 是人数，但是它的解却是 $x=1/2$. 此书的一整节都致力于解决初等的测量问题，我们将在后面讨论. 还有一个简短的章节用来写“三率法”，但它们中没有一个提供了二次方程式的实际运用. 最后，此书的后半部分完全用来处理继承遗产问题. 提出了数十个复杂的情况，而要解决它们必须要熟悉伊斯兰的遗产法. 实际所需的数学绝不会比解线性方程更复杂. 人们得到的结论只能是说，尽管花拉子米热衷于教他的读者如何解决数学问题，特别是如何处理二次方程式，但他没有能考虑任何需要这类方程的实际生活情景. 在这一方面，从巴比伦时代以来事情就很少有什么变化.

7.2.2 塔比·伊本·库拉和阿布·卡米勒的代数学

在花拉子米和伊本·吐克写了他们的著作后的 50 年间，伊斯兰数学家们认定二次方程式代数解必要的几何基础应该是欧几里得的著作而非古老的传统. 或许这些合理性论证中最早的是由塔比·伊本·库拉(Thābit ibn Qurra, 830—890)给出的. 塔比出生于哈兰(现土耳其南部)，被智慧宫的一位学者发现，约在 870 年被带到了巴格达，后来在那儿他成了一位伟大的学者. 在他所写的许多数学专著中有一本小小的书，书名是《论以几何证明验证代数问题》. 例如，对解方程 $x^2+bx=c$，塔比利用了图 7.8，其中 AB 代表 x，正方形 $ABCD$ 代表 x^2，BE 代表 b. 因此得到 $DE=AB\times EA$ 表示 c. 如果 W 是 BE 的中点，则欧几里得《原本》II-6 表明 $EA\times AB+BW^2=AW^2$. 由于 $EA\times AB$ 和 BW^2 已知(各等于 c 和 $\left(\frac{b}{2}\right)^2$)，从而 AW^2，进而 AW 便可求得，最后 $x=AB=AW-BW$ 便被确定. 塔比

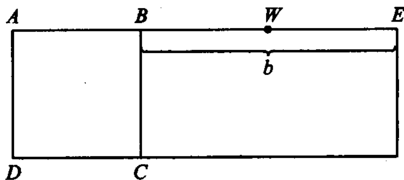


图 7.8 塔比·伊本·库拉对 $x^2+bx=c$ 解的几何论证.

很明显地了解到,《原本》II－6 的几何步骤与"代数学家"采取的步骤(即花拉子米所陈述的算法)完全是类似的,因而为这种算法提供了所必需的正确性论证. 塔比还指出如何用同样的命题去解 $x^2 = bx + c$,以及如何用《原本》II－5 去解 $x^2 + c = bx$.

埃及数学家阿布·卡米勒(Abū Kāmil ibn Aslam,850—930)在他著的《代数学》中也用《原本》II 给出了对这些解的相似推断:"我将用几何图形来解释它们的规则,而这些图形已被从事几何的智者们搞清楚了,并在欧几里得的书中有所解释."[17] 但是与塔比不同的是,阿布·卡米勒在他的讨论过程中重新证明了欧几里得的结果,并给出了数值的例子,实际上就是与花拉子米相同的那些最初的数值例子. 像其先辈那样,阿布·卡米勒在他用各种代数规则处理各种形式的二次方程的讨论之后,列出了大量的问题. 较之于早先的数学家而言他的进步之处在于他考虑了许多更为复杂的恒等式和复杂的问题,特别包括了对不尽根的处理.

阿布·卡米勒一点都不怵于处理"无理数". 在他的问题中他自由地使用它们. 像花拉子米的问题一样,他的许多问题也是以"将 10 分成两部分" 开始. 例如,考虑第 37 个问题:"如果人们说,10 已被分成了两部分,一部分被自乘而另一部分被乘以 8 的根,把乘以 8 的根的那个积从 …… 自乘的那部分的积中减去,得到了 40."[18] 这种情形的方程式为 $(10 - x)(10 - x) - x\sqrt{8} = 40$. 重写为形式 $x^2 + 60 = 20x + \sqrt{8x^2}\,(= (20 + \sqrt{8})x)$;阿布·卡米勒对平方和数等于根的情形执行了相应算法,最后得到 $x = 10 + \sqrt{2} - \sqrt{42 + \sqrt{800}}$,而另一部分 $10 - x$ 等于$\sqrt{42 + \sqrt{800}} - \sqrt{2}$.

阿布·卡米勒甚至应用替换法来简化问题并由此而处理了次数大于3的方程式,只要它们能化成二次形式就行. 第 45 问题能解释这两个想法:"有人说 10 被分成了两部分,每部分被另一部分去除,而且每个商式都作了自乘,然后从较大的一个中减去较小的那个则得到 2."[19] 这个方程是

$$\left(\frac{x}{10 - x}\right)^2 - \left(\frac{10 - x}{x}\right)^2 = 2.$$

阿布·卡米勒造了一个新的"东西"y,它等于$\frac{10 - x}{x}$ 并导出了新方程式$\frac{1}{y^2} = y^2 + 2$. 两端乘以 y^2 给了他一个关于 y^2 的二次方程:$(y^2)^2 + 2y^2 = 1$,其解为 $y^2 = \sqrt{2} - 1$. 因此 $y = \sqrt{\sqrt{2} - 1}$. 于是

$$\frac{10 - x}{x} = \sqrt{\sqrt{2} - 1},$$

然后阿布·卡米勒先将此方程两边平方再解 x. 最后结果是 $x = 10 + \sqrt{50} - \sqrt{50 + \sqrt{20\,000} - \sqrt{5000}}$.

在考虑阿布·卡米勒的代数时,请记住,像他那时的所有伊斯兰教科书那样,它是没有用任何符号来写的. 因此以现代的符号系统作出的几乎显然的代数处理在那里是完全用文字来进行的. 然而更为重要地,阿布·卡米勒乐于使用在花拉子米的时代已被系统化了的对任何类型正"数" 的代数算法. 他并不区分是用 2,或是$\sqrt{8}$,甚至是$\sqrt{\sqrt{2} - 1}$ 来进行运算. 由于这些算法是从几何中来的,就这个层面而言并不令人惊奇. 毕竟希腊人找不出正方形对角线的"数值" 表示,这正是他们使用线段和面积的几何代数的一条理由. 但是在处理这些量时,阿布·卡米勒对它们全体都以相同的方式进行诠释. 一个量从技术角度而言,它是一个平方,还是四次幂,是一个根式还是根的根这都没有关系. 对阿布·卡米勒来说,二次方程的解并不像在《原本》中某个命题所解释的那样;他认为它不是个线段. 它是一个"数",尽管阿布·卡米勒可能给不出这个词的真正定义. 因此他毫无顾忌地应用一般的规则把解中出现的各式各样的量结合在一起. 阿布·卡米勒以同一技巧来掌握所有这些量的意向有助于铺设出一条通向对数的概念有全新理解的道路,它与塞毛艾勒采用十进小数近似具

有同等的重要性.

7.2.3 凯拉吉,塞毛艾勒以及多项式代数

由花拉子米和阿布·卡米勒开始进行的把算术联系于代数的进程,因凯拉吉(Abū Bakr al-Karajī,1019)和塞毛艾勒的工作而在阿拉伯世界中一直延续了两个多世纪.在揭示出算术技巧可以富有成果地应用于代数方面,以及反之,在那些源于代数而发展起来的思想在处理数时的重要性方面,这后两位数学家都发挥了重大作用.

除了知道凯拉吉于1000年左右在巴格达工作并写了许多数学著作及工程专题方面著作外,对他其余的生活则没有什么了解.在11世纪的头十年里他写了一本主要的代数著作《发赫里》(al-Fakhrī,意思为"奇妙之事").依凯拉吉看来,《发赫里》的目标,这也是更一般的代数学的目标,是"从已知量出发确定未知量."[20] 他利用可以转化为处理未知量的全部算术技术来追求他的目标.他以系统地研究幂次代数为开端.尽管早先的著者们,其中包括丢番图,已经考虑过未知量的大于3次的幂,但是凯拉吉却是第一个充分了解这些幂次可以无限扩大的人.事实上,他阐述了一种对各次幂 x^n 和它们的逆 $\frac{1}{x^n}$ 命名的方法.每个幂可以由 x 乘前一次的幂递推地定义.由此得到了成比例的无限序列

$$1 : x = x : x^2 = x^2 : x^3 = \cdots,$$

及相似的倒数序列

$$\frac{1}{x} : \frac{1}{x^2} = \frac{1}{x^2} : \frac{1}{x^3} = \frac{1}{x^3} : \frac{1}{x^4} = \cdots.$$

一旦幂次被了解了,凯拉吉便建立了加、减及乘以单项式和多项式的一般步骤.然而在除法中他只用单项式去作除法因子,部分原因是他还不能把负数的规则结合进去,部分则由于他的以文字来表达的方式.类似地,虽然他发展出了一套计算多项式平方根的算法,但它却只能在一定的条件下使用.

凯拉吉继续了阿布·卡米勒的将算术运算应用于无理量方面的工作,并有较大的成功.特别,他明晰地诠释了在《原本》X中的各种不可公度类,把它们当作了"数"的类,并在它们上面定义了各种算术运算.像阿布·卡米勒一样,他并没有给出"数"的定义,只是用数值的而不是几何的技巧去对付各种不尽根式的量.作为这个进程的一部分,他发展出各种涉及不尽根的公式,例如

$$\sqrt{A+B} = \sqrt{\frac{A+\sqrt{A^2-B^2}}{2}} + \sqrt{\frac{A+\sqrt{A^2-B^2}}{2}}$$

及

$$\sqrt[3]{A} + \sqrt[3]{B} = \sqrt[3]{3\sqrt[3]{A^2B} + 3\sqrt[3]{AB^2} + A + B}.$$

处理代数运算方面进一步的工作是由塞毛艾勒完成的,他特别地引进了负系数.他把他处理这些系数的规则在他的代数教科书中表达得十分清楚,这本书的名称是《光辉的计算》(Al-Bābir fi'l-hisāb):

> 如果我们把一个加数从一个空幂中减去[$0x^n - ax^n$]则留下了相同的一个减数;如果我们从一个空幂中减去一个减数[$0x^n - (-ax^n)$]则留下的是那个相同的加数.如果我们从一个减数中减去一个加数,则留下的是他们减数的和;如果我们从一个较大的减数中减去一个较小的减数,结果是差的减数;如果被减数小于减数,则结果是它们差的加数.

人物小传	塞毛艾勒(1125—1174)(Al-Samaw'al)
	塞毛艾勒生于巴格达,其父母是具良好教育的犹太人.他的父亲实际上是位希伯来语的诗人.除去给予他宗教教育外,他的父母鼓励他学习医学和数学.由于那时在巴格达不再有智慧宫,他不得不独立地学习数学并在中东各处旅行.在他只有十九岁时便写成了他的主要著作《光辉的计算》.之后,他的兴趣转向了医学,并成为了一个成功的医生及医学教科书的作者.惟一尚存的一本书的标题为《伴侣在爱园的漫步》,这是关于性医学的专著以及色情故事的集子.在他大约40岁时,他决定皈依伊斯兰教.为了向世界表明他的皈依的正确性,在1167年他写了一本自传,陈述他反对犹太教的辩言,而这部著作则成了伊斯兰教辩论学中反犹太倾向的一个来源.

给出了这些规则后,塞毛艾勒便能容易地以合并同类项的办法进行多项式的加和减.当然,对乘法他需要指数的规律.本质上说,凯拉吉以及阿布·卡米勒和其他一些人都曾用过这个规律.但是,就像例子所表明的那样,将一个平方和一个立方的乘法用语言表达成平方-立方,不能看出指数相加的数值性质.塞毛艾勒决定最好能用由各列构成的表格来表达这个规律,其中每列代表了一个数或一个未知量的不同幂.事实上,他也明白他可以像处理 x 的幂那样容易地来处理 $1/x$ 的幂.在他的表格中,各列之首放的是数值,从中间的标以0的那列向两边读过去.例如,以2开头的左边一列被称作"平方"而左边以5开头的一列称作"平方-立方",右边以3开头的那列称作"立方的一份",等等.为了简化起见,我们只使用 x 的幂.在他对其规则的最初解释中,塞毛艾勒还在左端1列的下方放了一个特定的数,如像2,然后在各相应列下放上2的各次幂:

7	6	5	4	3	2	1	0	1	2	3	4	5	6	7
x^7	x^6	x^5	x^4	x^3	x^2	x	1	x^{-1}	x^{-2}	x^{-3}	x^{-4}	x^{-5}	x^{-6}	x^{-7}
128	64	32	16	8	4	2	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{128}$

现在塞毛艾勒便利用这个图表来解释我们所称的指数法则,$x^m x^n = x^{m+n}$:"乘积的阶到其中一个因子的阶的距离等于另一个因子的阶到单位阶的距离.如果这两个因子在不同的方向,则我们计算离第一个因子的阶的距离是朝向单位阶方向,如果它们处在同一方向,我们则朝离开单位阶来计算."[22] 故而,举例来说,x^3 乘以 x^4 时,朝3列的左边数4阶便得到结果 x^7.对 x^3 乘以 x^{-2},则从第3列朝右数2阶得答案 x^1.运用这些规则塞毛艾勒便能容易地进行 x 与 $\frac{1}{x}$ 的多项式乘法以及多项式除以单项式.

塞毛艾勒用一个类似的图表也可以作一个多项式除以另一个多项式.这个新图表使我们想起了用以解多项式方程的中国算板.这个图表的每列仍代表 x 或 $\frac{1}{x}$ 的给定幂.但是现在每列中的数代表的是除法过程中涉及到的多项式的系数.例如用 $6x^2 + 12$ 去除 $20x^2 + 30x$,他先把20和30分别放在以 x^2 和 x 开头的各列里,并把6和12放在以 x^2 和1开头的列中的下方.由于在 x 列的除法因子中有一个"空阶",他在那里放了0.第二步他用 $6x^2$ 去除 $20x^2$,得到 $3\frac{1}{3}$.把这个数放在单位列的对应于答案的那一行的位置上.$3\frac{1}{3}$ 乘以 $6x^2 + 12$ 是 $20x^2 + 40$.下一步是减法.在 x^2 列上剩下的自然为0.在 x 列剩下的是30,而单位列中剩下的是 -40.现在塞毛艾勒写下一张新的图表,其中6,0,12向右边移动了一位,给出的进一步指令是将它去除 $30x - 40$.$30x$ 除以 $6x^2$ 的头一次商是 $5\cdot\frac{1}{x}$,从而将其放在了以 $\frac{1}{x}$ 开头的那一列的答案位置,步骤这样继续下去.这里展示的是塞毛艾勒做他的除法

问题时的头两张图表：

x^2	x	1	$\frac{1}{x}$	$\frac{1}{x^2}$	$\frac{1}{x^3}$
		$3\frac{1}{3}$			
20	30				
6	0	12			

x^2	x	1	$\frac{1}{x}$	$\frac{1}{x^2}$	$\frac{1}{x^3}$
		$3\frac{1}{3}$	5		
	30	-40			
	6	0	12		

在这个特定的例子中的除法不是完全除尽的.通过八步后塞毛艾勒得到：

$$3\frac{1}{3}+5\left(\frac{1}{x}\right)-6\frac{2}{3}\left(\frac{1}{x^2}\right)-10\left(\frac{1}{x^3}\right)+13\frac{1}{3}\left(\frac{1}{x^4}\right)$$
$$+20\left(\frac{1}{x^5}\right)-26\frac{2}{3}\left(\frac{1}{x^6}\right)-40\left(\frac{1}{x^7}\right).$$

为展现他的乘法步骤的流畅自如,现在他用乘数去乘它以检验答案.因为这个积与被除因子只在$\frac{1}{x^6}$和$\frac{1}{x^7}$项有差别,所以他称所得结果为“近似答案”.虽然如此,他也注意到商中系数间存在的一种模式.事实上,如果 a_n 表示$\frac{1}{x^n}$的系数,则此模式由 $a_{n+2}=-2a_n$ 给出.于是他不无自豪地写出这个商的后面的 21 项,其最后一项为 $54\,613\frac{1}{3}\left(\frac{1}{x^{28}}\right)$.

在知道了塞毛艾勒把多项式的除法拓展到$\frac{1}{x}$的多项式上去,并想成是作为近似的部分结果之后,那么,对他会简单地用 10 代替 x 去作整数的除法就不会觉得惊奇了.正如已经注意的那样,塞毛艾勒是第一个明确认识到人们可以用计算越来越多的小数位数来逼近一个分数的人.在发展代数处理与数的处理是平行的这种思想中,凯拉吉和塞毛艾勒发挥了非常重要的作用.实际上,任何应用于其中之一的技巧也可以适用于另外一个.

7.2.4 归纳、幂的和以及帕斯卡三角形

先由凯拉吉后由塞毛艾勒和其他人引进的另一个重要思想是在处理某些算术序列时使用的归纳推理.凯拉吉使用这种推理证明了整数立方和的一个已为阿耶波多所知道的结果(或许甚至已为希腊人所知).但是凯拉吉并未对任意 n 给出一般性结果.他只对特别的整数 10 叙述了定理：

$$1^3+2^3+3^3+\cdots+10^3=(1+2+3+\cdots+10)^2.$$

然而他的证明方法很清楚表明是能推广到任意其他整数的.

考虑具边长 $1+2+3+\cdots+10$ 的正方形 $ABCD$(图 7.9).令 $BB'=DD'=10$,并作出磬折形 $BCDD'C'B'$,凯拉吉算出此磬折形的面积为

$$2\cdot 10(1+2+\cdots+9)+10^2=2\cdot 10\cdot\frac{9\cdot 10}{2}+10^2$$
$$=9\cdot 10^2+10^2=10^3.$$

因为正方形 $ABCD$ 的面积等于正方形 $AB'C'D'$ 的面积与此磬折形面积的和,从而得到 $(1+2+\cdots+10)^2=(1+2+\cdots+9)^2+10^3$. 按此方式继续下去直到最后面积为 $1=1^3$ 的正方形 $A\hat{B}\hat{C}\hat{D}$,凯拉吉便由正方形 $ABCD$ 的面积等于正方形 $A\hat{B}\hat{C}\hat{D}$ 的面积与各个磬折形的面积 $2^3,3^3,\cdots,10^3$ 之和证明了他的定理.

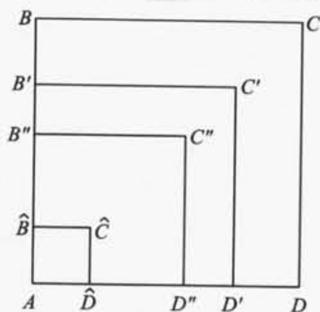


图 7.9 对整立方和公式的凯拉吉的证明.

凯拉吉的推理实质上包含了现代归纳推理的两个部分,即论断对 $n=1(1=1^3)$ 成立,及由 $n=k-1$ 成立推出 $n=k$ 成立. 当然第二步并不清晰,因为从某种意义上说,凯拉吉的推理是反向的,就是说,他由 $n=10$ 开始向下到1而不是向上进行. 不管怎样,他在《法赫里》中的这个推理是现存的对整立方和公式最早的证明.

整数和与它们的平方和的公式久为人知,而对立方和的公式只要人们考虑几个例子也容易发现. 然而要给出一个能推广到四次幂和公式的推理却是比较困难的,但它却在11世纪初由埃及数学家伊本·海塞姆(965—1039)在一本著作中做到了. 他没有把他的结果推广到求更高次幂的和大概因为在他计算抛物体体积时只需要二次和四次幂和的公式的缘故吧. 他的这个工作将在后面的7.4.4节中讨论.[23]

人物小传	伊本·海塞姆(965—1039)(Ibn Al-Haytham)
	伊本·海塞姆(图7.10)在欧洲以阿尔哈琛(Alhazen)这个名字而知名,他被认为是最具影响的伊斯兰科学家之一. 他生于巴什拉,现在位于伊拉克,但在被哈奇姆哈里发邀请去参与尼罗河控制工程工作后便在埃及度过了他的大半生. 虽然工程计划没有任何成果,但是伊本·海塞姆却在埃及写出了他最主要的著作,共有七册的《光学》. 在13世纪初,《光学》被翻译成拉丁文. 从此之后的许多世纪中,此书在欧洲得到研读和评论. 伊本·海塞姆作为一个数学家的名声主要在于他对所谓"阿尔哈琛问题"的处理,这些问题是关于在某些反射面上找出一个或一些点,使得光线从曲面外两点中的一点经过这一(些)点恰好反射到另一点. 在《光学》的第五册中他努力对各种曲面,球面,柱面,圆锥面,凸的或凹的情形来解决这个问题. 虽然他并未取得完全成功,但他的成就表明他充分掌握初等的和先进的希腊的几何学. 在他生命的最后年代里,伊本·哈塞姆以复制书籍为生,这些著作中包括了欧几里得的《原本》,阿波罗尼奥斯的《圆锥曲线论》和托勒密的《天文学大成》.

在证明求和公式中,伊本·海塞姆的中心思想是推导出等式

$$(n+1)\sum_{i=1}^{n}i^k=\sum_{i=1}^{n}i^{k+1}+\sum_{p=1}^{n}\left(\sum_{i=1}^{p}i^k\right).\qquad(7.2)$$

伊本·海塞姆没有以这个结果的一般形式来叙述而是只对特别的整数 $n=4$ 及 $k=1,2,3$ 来叙述的. 然而,像凯拉吉使用归纳推理一样,他的证明立刻可以推广到 n 和 k 的任意值. 我们来考察他对 $k=3$ 和 $n=4$ 时的证明:



图 7.10 为纪念伊本·海塞姆在光学上的工作而发行的巴基斯坦邮票.

$$(4+1)(1^3+2^3+3^3+4^3)=4(1^3+2^3+3^3+4^3)+1^3+2^3+3^3+4^3$$

$$= 4 \cdot 4^3 + 4(1^3 + 2^3 + 3^3) + 1^3 + 2^3 + 3^3 + 4^3$$
$$= 4^4 + (3 + 1)(1^3 + 2^3 + 3^3) + 1^3 + 2^3 + 3^3 + 4^3.$$

但是已假定等式(7.2)对 $n = 3$ 时成立,故有

$$(3 + 1)(1^3 + 2^3 + 3^3) = 1^4 + 2^4 + 3^4 + (1^3 + 2^3 + 3^3) + (1^3 + 2^3) + 1^3.$$

因此等式(7.2)对 $n = 4$ 成立.把这个证明推理重写为对 n 归纳的现代证明是直接了当的事.

伊本·海塞姆利用等式(7.2)来推导整数幂和公式,而且公式是以其一般形式叙述的.这样,对 $k = 2$ 和 $k = 3$ 我们有

$$\sum_{i=1}^{n} i^2 = \left(\frac{n}{3} + \frac{1}{3}\right)n\left(n + \frac{1}{2}\right) = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6},$$

$$\sum_{i=1}^{n} i^3 = \left(\frac{n}{4} + \frac{1}{4}\right)n(n + 1)n = \frac{n^4}{4} + \frac{n^3}{2} + \frac{n^2}{4}.$$

我们在这里不去考虑这些结果的证明而只考虑对四次幂的类比结果的证明.虽然是以完全一般形式(在最后部分)叙述的,这个结果也只对 $n = 4$ 的情形作了证明.然而,我们可以把这个证明看成为"可推广性的例子"方法的一个代表,这个方法早在欧几里得的工作中已能看到.不管怎样,伊本·海塞姆把关于三次和二次的公式代入了(7.2)式,并用它们来证明了对四次幂的公式:

$$\begin{aligned}(1^3 + 2^3 + 3^3 + 4^3)5 &= 1^4 + 2^4 + 3^4 + 4^4 + (1^3 + 2^3 + 3^3 + 4^3) \\ &\quad + (1^3 + 2^3 + 3^3) + (1^3 + 2^3) + 1^3 \\ &= 1^4 + 2^4 + 3^4 + 4^4 + \left(\frac{4^4}{4} + \frac{4^3}{2} + \frac{4^2}{4}\right) + \left(\frac{3^4}{4} + \frac{3^3}{2} + \frac{3^2}{4}\right) \\ &\quad + \left(\frac{2^4}{4} + \frac{2^3}{2} + \frac{2^2}{4}\right) + \left(\frac{1^4}{4} + \frac{1^3}{2} + \frac{1^2}{4}\right) \\ &= 1^4 + 2^4 + 3^4 + 4^4 + \frac{1}{4}(1^4 + 2^4 + 3^4 + 4^4) \\ &\quad + \frac{1}{2}(1^3 + 2^3 + 3^3 + 4^3) + \frac{1}{4}(1^2 + 2^2 + 3^2 + 4^2) \\ &= \frac{5}{4}(1^4 + 2^4 + 3^4 + 4^4) + \frac{1}{2}(1^3 + 2^3 + 3^3 + 4^3) \\ &\quad + \frac{1}{4}(1^2 + 2^2 + 3^2 + 4^2),\end{aligned}$$

$$\begin{aligned}1^4 + 2^4 + 3^4 + 4^4 &= \frac{4}{5}(1^3 + 2^3 + 3^3 + 4^3)\left(4 + \frac{1}{2}\right) - \frac{1}{5}(1^2 + 2^2 + 3^2 + 4^2) \\ &= \frac{4}{5}\left(4 + \frac{1}{2}\right)\left(\frac{4}{4} + \frac{1}{4}\right)4(4 + 1)4 - \frac{1}{5}\left(\frac{4}{3} + \frac{1}{3}\right)4\left(4 + \frac{1}{2}\right),\end{aligned}$$

最后得

$$1^4 + 2^4 + 3^4 + 4^4 = \left(\frac{4}{5} + \frac{1}{5}\right)4\left(4 + \frac{1}{2}\right)\left[(4 + 1)4 - \frac{1}{3}\right].$$

由对 $n = 4$ 的这个结果,伊本·海塞姆便径直用文字说出了他的一般性结果;我们可以把它翻译成现代公式,即

$$\sum_{i=1}^{n} i^4 = \left(\frac{n}{5} + \frac{1}{5}\right)n\left(n + \frac{1}{2}\right)\left[(n + 1)n - \frac{1}{3}\right].$$

这个时期的另一个归纳推理是关于二项式定理和帕斯卡三角形的,这可在塞毛艾勒的书《光辉的计算》中找到,他在书中提到了凯拉吉对这方面课题的处理.但是,由于凯拉吉有关这方面的著

作业已佚失,我们来考虑塞毛艾勒的说法.二项式定理是

$$(a+b)^n=\sum_{k=0}^{n} C_k^n a^{n-k} b^k,$$

其中 n 为正整数,C_k^n 为二项式系数,即帕斯卡三角形中的各个分量.自然,塞毛艾勒并没有用符号系统来表示它们,而是对每个个别的例子用文字来写出这个公式.例如,在 $n=4$ 的情形,他写道:"对一个被分成两部分的数,它的平方-平方[即四次幂]等于每个部分的平方-平方,每一部分与另一部分立方乘积的四倍,及每一部分平方乘积的六倍."[24] 塞毛艾勒提供了一张二项式系数的表来表明该如何将此规律推广到较大的 n 值:

x	x^2	x^3	x^4	x^5	x^6	x^7	x^8	x^9	x^{10}	x^{11}	x^{12}
1	1	1	1	1	1	1	1	1	1	1	1
1	2	3	4	5	6	7	8	9	10	11	12
	1	3	6	10	15	21	28	36	45	55	66
		1	4	10	20	35	56	84	120	165	220
			1	5	15	35	70	126	210	330	495
				1	6	21	56	126	252	462	792
					1	7	28	84	210	462	924
						1	8	36	120	330	792
							1	9	45	165	495
								1	10	55	220
									1	11	66
										1	12
											1

他提供的构造此表的步骤是我们所熟悉的,其中任何一个分量由它左边那一个分量加这分量正上方的那个分量得到.然后他指出人们可以利用这张表读出"一个分成两部分的数"的直到十二次的任意幂次的展开式.

记住这张表后让我们来看一下塞毛艾勒是如何证明上面所引述的 $n=4$ 的结果的.假定数 c 等于 $a+b$,因为 $c^4=c\cdot c^3$,而 c^3 已知由 $c^3=(a+b)^3=a^3+b^3+3ab^2+3a^2b$ 给出,从而得到 $(a+b)^4=(a+b)(a+b)^3=(a+b)(a^3+b^3+3ab^2+3a^2b)$.反复使用结果 $(r+s)t=rs+rt$,这是塞毛艾勒从欧几里得的《原本》II 中引用的,他发现这最后的量等于

$$\begin{aligned}&(a+b)a^3+(a+b)b^3+(a+b)3ab^2+(a+b)3a^2b\\&=a^4+a^3b+ab^3+b^4+3a^2b^2+3ab^3+3a^3b+3a^2b^2\\&=a^4+b^4+4ab^3+4a^3b+6a^2b^2.\end{aligned}$$

这里的系数都是表中适当的分量,而且这个展开式表明了新的系数恰恰是按在构造表时所说的那样,由老的展式的系数形成的.塞毛艾勒下一步引述了 $n=5$ 的结果,并对一般性结果下了断言:"凡是理解了我们刚才所说的内容的人,就能对任何分成两部分的数证明:它的平方-立方[5 次幂]等于它的各部分平方-立方的和,以及和它每部分与另一部分的平方-平方之积的 5 倍,每部分平方与另部分立方之积的 10 倍.对各个升高了的幂次也如此."[25] 像凯拉吉和伊本·海塞姆一样,塞毛艾勒的推理包含了归纳证明的两个基本要素.他由已知结果的值出发,这里是 $n=2$,然后对给定的整数的结果去推出对下一个数的结果.尽管塞毛艾勒没有任何办法来表述并证明一般的二项式定理,但对现代的读者而言,只要假定在定理的陈述中其系数本身也是归纳定义的,即本质上像塞毛

艾勒那样定义为 $C_m^n = C_{m-1}^{n-1} + C_m^{n-1}$,那么从塞毛艾勒的推理到二项式定理的完全归纳证明只不过是小小的一步罢了.不管什么情形,帕斯卡三角形在伊斯兰世界,也像我们知道的那样在中国,都被用来发展一种计算数的根的算法.在伊斯兰这方面,有文献纪录的这种算法是从塞毛艾勒那时开始的,然而有很强的迹象表明至少早在一个世纪之前就已经知道它了.

7.2.5 奥玛·海亚姆和三次方程式的解

伊斯兰世界的代数学发展,除了它的算术化和它的归纳思想的发展外还有另一个组成部分,即几何学的应用.9 世纪末,伊斯兰数学家在读了主要的希腊著作后已经注意到某些几何问题可导致三次方程,而这些方程式又可以通过求两条圆锥截线的交得到解.这些问题中包括了倍立方体以及阿基米德的将球分为两块使其体积具已知比.在 10 世纪和 11 世纪中,许多伊斯兰数学家也曾以这种圆锥曲线相交的希腊人的想法解过某些三次方程.但是,正是数学家兼诗人的乌马尔·伊本·易卜拉辛·海亚米(Umar ibn Ibrāhīm al-Khayyāmī,1048—1131)(在西方通常被称为奥玛·海亚姆(Omar Khayyam)),他首先以这种一般性的方法对三次方程式进行了系统的分类并解出了所有类型的这种方程.

奥玛·海亚姆的主要数学著作《论代数问题的证明》主要从事于解三次方程.正如作者在他的序言中清楚说明的那样,这部著作的读者必须完全熟悉欧几里得的《原本》及《数据》("Data"),还有阿波罗尼奥斯的《论圆锥曲线》的头两册,因为三次方程只能用圆锥曲线的性质来几何地解出.但是此书处理的是代数的而不是几何的问题,而且海亚米宁愿提出解三次方程的代数算法,就像花拉子米解二次方程的那三种算法.正如他写的,"但是当问题的对象是个纯粹的数时,既非我们也非那些关心代数的人们中任何一个人有能力解此方程 —— 或许将来那些追随我们的其他人能够填补这个空隙."[26] 直到了 16 世纪"乌马尔"的希望才得以实现.

人物小传	奥玛·海亚姆(1048—1131)(Al-Khayyāmī)
	海亚姆出生于伊朗的尼沙普尔,那是在塞尔柱突厥人征服了这片土地的 1048 年.在他生涯的大部分时间中他享有塞尔柱统治者们的支持.事实上,他作为一个历法改进小组的领头人,在伊斯法汗的观测站度过了许多年.当统治者权位更迭时他曾多次失宠,但最终总能聚集到足够的资助以写出许多数学和天文学著作,还有诗歌和哲学的著作.事实上,他在西方世界是以他的诗集《鲁拜集》而最为闻名.在他的伟大的代数著作的序言中,他抱怨他的工作曾是多么困难,随后则感谢统治者给予他所需的支持: 长久以来,我被令人心烦的障碍所阻挡,我找不出时间来完成这本著作,或者根本不能把思绪集中在它上面.…… 我们大部分的同代人都是些伪科学家,他们把真理和虚假混为一谈,他们不以欺诈和卖弄学问为耻,他们把所知的那一点科学知识只用于低贱的物欲目的.当他们看到一个渴望追寻真理的卓越人物,看到一个宁要诚实并尽力排斥虚假和谎言、回避虚伪和背信的人时,他们便侮辱他,捉弄他.在我绝望之际,蒙真主垂恩,使我遇见了我们举世无双的君王、至高无上的法官,赛义德·阿布·塔希尔依玛目阁下 …… 一个集科学中无上的权力和行动的坚定性于一身的人,并赐予我他的亲密友谊 …… 我的心因见到他而极度欣喜 …… 他的开明和宠爱使我的力量得到增强.为了能够更接近他的崇高地位,我觉得我自己有义务重新捡起我的工作,以总结我已证实了的哲学理论中的精髓,这些工作曾因纷乱多变的时局而被我放弃过.[27]

按花拉子米的风格,海亚米以对不超过三次的方程进行完全分类来开始他的工作.由于对海亚

米来说,像他的先辈那样,所有的数都是正的,所以他不得不把具有正根的各种形式分列出来.其中有十四个是不能简约为二次或线性方程式的.一共有三组,一组是二项的方程式 $x^3 = d$,另一组是三项方程式,$x^3 + cx = d, x^3 + d = cx, x^3 = cx + d, x^3 + bx^2 = d, x^3 + d = bx^2$,和 $x^3 = bx^2 + d$;最后是四项方程式,$x^3 + bx^2 + cx = d, x^3 + bx^2 + d = cx, x^3 + cx + d = bx^2, x^3 = bx^2 + cx + d$, $x^3 + bx^2 = cx + d, x^3 + cx = bx^2 + d$ 和 $x^3 + d = bx^2 + cx$.作者对它们中每个方程都作了仔细分析.他描述了求解方程所需要的圆锥曲线,并证明了他的解是正确的.最后还讨论了没有解或有多于一个解时的条件.我们将在这里讨论海亚米对 $x^3 + cx = d$ 的解,或者像他所说的,是"一个立方及边等于一个数"的情形.

与凯拉吉和塞毛艾勒不一样,海亚米小心翼翼地保持了希腊人的齐性思想,就是说,他把三次方程设想为立体之间的一个方程式,由于 x 表示了立方体的一条边,c 必定代表了一块面积(可以表示成一个正方形).故而 cx 是个立体,而 d 自身也表示了一个立体.为构造这个解,海亚米令 AB 等于正方形 c 的边长,或是说 $AB = \sqrt{c}$(图 7.11).然后他作 BC 垂直于 AB,使得 $BC \cdot AB^2 = d$ 或 $BC = d/c$.下一步,他按 Z 的方向延长 AB 并作一抛物线,它以 B 为顶点,BZ 为轴,AB 为参数.按现代的记号,这个抛物线的方程为 $x^2 = \sqrt{c}y$.类似地,他在直线 BC 上作一个半圆.它的方程式是

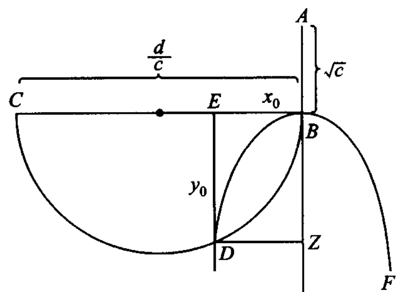


图 7.11　海亚米对 $x^3 + cx = d$ 的解的构造.

$$\left(x - \frac{d}{2c}\right)^2 + y^2 = \left(\frac{d}{2c}\right)^2$$

或

$$x\left(\frac{d}{c} - x\right) = y^2.$$

这个圆与抛物线交于点 D.正是这个点的 x 坐标,即这里以线段 BE 表示的,给出了此方程式的解.

利用抛物线与此圆的基本性质海亚米证明了他的解是正确的.如果 $BE = DZ = x_0, BZ = ED = y_0$,则首先有 $x_0^2 = \sqrt{c}y_0$ 或 $\frac{\sqrt{c}}{x_0} = \frac{x_0}{y_0}$,这是因为 D 在抛物线上,其次因为 D 在此半圆上,有 $x_0\left(\frac{d}{c} - x_0\right) = y_0^2$ 或 $\frac{x_0}{y_0} = \frac{y_0}{\frac{d}{c} - x_0}$.由此得到

$$\frac{c}{x_0^2} = \frac{x_0^2}{y_0^2} = \frac{y_0^2}{\left(\frac{d}{c} - x_0\right)^2} = \frac{y_0}{\frac{d}{c} - x_0} \cdot \frac{x_0}{y_0} = \frac{x_0}{\frac{d}{c} - x_0},$$

于是 $x_0^3 = d - cx_0$,即 x_0 是所要的解.在没有展示任何证明的情形下,海亚米指出在这里这一类方程总有一个单一的解,换句话说,这条抛物线与圆除原点外总交于一个点.原点却不是问题的解.海亚米的注解所反映的是一段现代的论述,即方程 $x^3 + cx = d$ 恰好有一个正的解.

海亚米以同样的方式处理了这十四种情形中的每一个.在那些并不总存在正解的情形,他对存在性给出了一个几何的条件.即是说,存在 0 个,一个或两个解的情形依赖于所涉及到的圆锥曲线是不相交还是交于一点还是交于两个点.在这个分析中他的一个错误是在方程 $x^3 + cx = bx^2 + d$ 的情形,这时他没有发现具有三个解的可能性.然而,总的来说,他没有把具有一个或两个解的存在性问题与系数满足的条件联系起来.即便在 $x^3 + d = bx^2$ 的情形时他做了点联系,也只是一个局限

的方式.在那个方程的情形,他指出如果$\sqrt[3]{d} = b$则方程没有解.因为如果x是个解,则$x^3 + b^3 = bx^2$,故$bx^2 > b^3$及$x > b$.由于$x^3 < bx^2$,从而又有$x < b$,引出矛盾.相似地,当$\sqrt[3]{d} > b$时也没有解.但是条件$\sqrt[3]{d} < b$却不能保证有一个解.海亚米还指出,按此问题中圆锥曲线(一条抛物线和一条双曲线)相交的次数来决定它有零个,一个或两个(正的)解.

7.2.6 萨拉夫·丁·图西和三次方程

海亚米的方法得到了萨拉夫·丁·图西(Sharaf al-Dīn al-Tūsī,1213)的改进.萨拉夫·丁是一位数学家,生于波斯的图斯.像他的先辈那样,他也以将三次方程分成若干组作为开始.他的分类不同于海亚米,因为他所感兴趣的是用以确定解的个数的加在系数上的条件.因此他的第一组是由那些可以简约为二次方程的方程和方程$x^3 = d$组成.第二组由八个三次方程组成,它们都至少有一个(正的)解.第三组是由那些可能有也可能没有(正)解的方程类型组成,它们是否有解取决于系数的特定值.这些包括了$x^3 + d = bx^2$,$x^3 + d = cx$,$x^3 + bx^2 + d = cx$,$x^3 + cx + d = bx^2$及$x^3 + d = bx^2 + cx$.

对第二组方程式他的解法与海亚米的一样.他以两个适当选取的圆锥曲线的交来确定方程的解.但他高明于海亚米之处在于他总是仔细描述出两条圆锥曲线为什么相交的道理.但第三组才是他作出了最具原创性贡献的地方.

考虑萨拉夫·丁对$x^3 + d = bx^2$的分析.他先将方程做成$x^2(b - x) = d$的形式.然后他指出此方程是否有解的问题取决于"函数"$f(x) = x^2(b - x)$是否达到值d.换句话说,他需要考虑$x^2(b - x)$的极大值问题(图7.12).他然后宣称值$x_0 = \frac{2b}{3}$事实上给出$f(x)$的极大值,即当x介于0与b之间时$x^2(b - x) \leqslant \left(\frac{2b}{3}\right)^2\left(\frac{b}{3}\right) = \frac{4b^3}{27}$.令人好奇的是他并没有说出他为什么选了$x_0$的这个特别的值.或许是类比于已为希腊人所知的事实(《原本》VI - 28)即$x = \frac{b}{2}$为$x(b - x)$的极大值而猜出来的,或是由于仔细研读了阿基米德的《论球面和柱面》II的问题4即一个涉及这种类型的三次方程的问题而知道的.也有人提示说,他是考虑了满足对$y < x$和$y > x$都有$f(x) - f(y) > 0$的x的条件而得到的,就是说,本质上是计算了$f(x)$的"导数"的一个零点.[28] 不管他是怎样推导出来的,他确实给了一个完全正确的几何证明,证明此值确为极大.对他的第三组中五个方程式中的每一个他都能完成类似的分析.

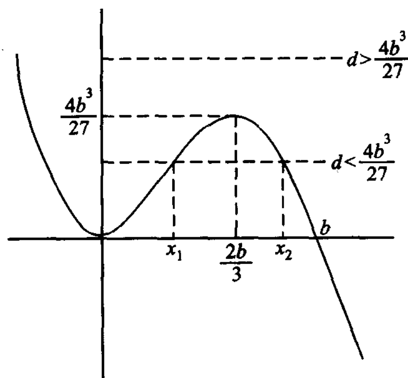


图7.12 对三次方程$x^3 + d = bx^2$的萨拉夫·丁·图西解释的现代图解.

知道在$\frac{2b}{3}$给出极大值后,萨拉夫·丁指出,如果极大值$4b^3/27$小于已知的d,则此方程无解.如果$\frac{4b^3}{27}$等于d,则只有一个解$x = \frac{2b}{3}$.最后,如果$\frac{4b^3}{27}$大于d,则有两个解x_1和x_2,其中$0 < x_1 < \frac{2b}{3}$,$\frac{2b}{3} < x_2 < b$.知道了解的存在性,萨拉夫·丁便将解此方程化为解另一个已知形式的方程式.在此情形下它化成了方程式$x^3 + bx^2 = k$,其中$k = \frac{4b^3}{27} - d$.他证明,如果以两条圆锥曲线的相交几何

地给出此方程的解为 X,那较大的那个解 x_2 由 $x_2 = X + \frac{2b}{3}$ 给出.为求出剩下的那个根 x_1,作者提供了一个新的解法.他先求出二次方程式 $x^2 + (b - x_2)x = x_2(b - x_2)$ 的解 Y,然后仍用几何的方式证明 $x_1 = Y + b - x_2$ 是原方程的另一个正根.于是这个新多项式的根通过这个变量的变换与老多项式的根联系了起来.显然由此可知,萨拉夫・丁对于三次方程的性质以及它们的根和系数间的关系有了扎实的了解.不同于前人之处是他能够看到各种类型的三次方程是相互关联的.但还要指出,虽然他实际上运用了一个三次方程的判别式,即这里的 $\frac{4b^3}{27} - d$ 来决定是否存在正根,但他还不能用它来代数地给出数值解.

另一方面,萨拉夫・丁对求这些三次方程的数值解颇有兴趣.他给出的讨论数值解的例子是 $x^3 + 14\,837\,904 = 465x^2$.用上面提到的方法,他首先计算出 $\frac{4b^3}{27} = 14\,895\,500$ 及 $k = \frac{4b^3}{27} - d = 57\,596$.由此知有两个解 x_1, x_2,满足 $0 < x_1 < 310$ 及 $310 < x_2 < 465$.为了求出 x_2,他需要解方程式 $x^3 + 465x^2 = 57\,596$.他求出了它的一个解为 11,因而 $x_2 = \frac{2b}{3} + 11 = 310 + 11 = 321$.为求 x_1 他需解二次方程式 $x^2 + 144x = 46\,224$.它的正解是个无理数,近似地等于 154.73,他是用与在第 6 章中讨论过的中国方法有关的数值法求得的.原方程的解 x_1 便等于 298.73.

伊斯兰的代数学家们显然在发展由他们从巴比伦人那里接收来的代数学方面作出了巨大的进展.他们把从希腊人得到的证明概念溶进了他们的工作中,因而将他们的方法置于一个坚实的根基之上.这个工作中的一些部分传到了欧洲,让欧洲人第一次看到了代数,同时,另一些部分可惜未能传过去,这迫使欧洲人不得不自己去重新发现某些思想.后面我们会讨论到这些思想的传播途径.

7.3　组合数学

我们已经看到,在 9 世纪或许更早的时候,印度人已经知道了组合和置换的基本公式.伊斯兰数学家也对这些研究有兴趣.例如,卡利勒・伊本・艾哈默德(717—791),一位词典编纂者,他的兴趣是将阿拉伯语言中的字进行分类,他计算了由阿拉伯文的 28 个字母中取 2,3,4 或 5 个字母组成的字的个数.另外,塞毛艾勒在他的《光辉的计算》中在讨论解大方程组的方法时,实际上写出了在 10 个未知量中每次取 6 个的全部 210 种组合,而且是用一种系统的方式写出来的.但是他并没有说明对其他情形如何计算这个数目.只有在 13 世纪,我们才看到证据表明已推导出了基本的组合公式.我们将考察许多伊斯兰数学家对此工作的贡献.

7.3.1　数颜色

早在 13 世纪,艾哈默德・阿布达瑞・伊本・穆恩依姆便讨论过从一个 n 元的集合中取 r 件东西的组合数,他是用取 $r - 1$ 件东西的组合来考察推导这个数的.关于伊本・穆恩依姆知道得很少,但知道他大概生活在穆罕默德・伊本・雅库伯・纳西尔(1199—1213)统治时期的马拉喀什(现于摩洛哥)的艾尔莫哈德宫廷里.虽然艾尔莫哈德王朝原本统治一大片包括了北非和西班牙的许多的地方,但是在 1212 年,一个基督教国王的联盟在西班牙拉斯纳瓦斯的一次战役中打败了纳西尔之后,便失去了他的西班牙领地的许多地方.

伊本・穆恩依姆(Ibn Mun'im) 基本上是在检验那个老问题,即从阿拉伯文字母中可能形成的

词的个数.但在处理这个问题之前,他考虑了一个不同的问题:由十种不同颜色的丝稠可以做成多少种不同颜色的丝绸捆?他做了仔细的计算.首先,他指出对一种颜色的捆有十种可能性,即 $C_1^{10}=10$.为了计算对两种色的可能性,伊本·穆恩依姆依次列出下面的偶对(其中 c_i 表示第 i 个颜色):

$$(c_2,c_1);(c_3,c_1),(c_3,c_2);\cdots(c_{10},c_1),(c_{10},c_2),\cdots,(c_{10},c_9)$$

并指出

$$C_2^{10}=C_1^1+C_1^2+\cdots+C_1^9=1+2+\cdots+9=45.$$

对 k 小于 10 的 C_2^k 可用相似的办法计算出.对 C_3^{10} 的计算,伊本·穆恩依姆按类似的方式进行.

> 为决定三种颜色的组合数,人们首先将第三种颜色组合上第一种和第二种,然后将第四种去组合前三种颜色所构成的每个偶对,这里的前三种是指第一,二,三种颜色.然后将第五种去组合前四种颜色所构成的每个偶对,…… 将第十种颜色去组合前九种颜色所构成的每个偶对.但是每个颜色偶对是由第二行取出的一个组合.[29]

换句话说,对每个 $c_k,k=3,4,\cdots,10$,穆恩依姆考虑所有指标小于 k 的前面计算过的偶对,例如

$$(c_3,(c_2,c_1));(c_4,(c_2,c_1)),(c_4,(c_3,c_1)),(c_4,(c_3,c_2));(c_5,(c_2,c_1)),\cdots.$$

因此 C_3^{10} 是和数 $1+3+6+\cdots+36=C_2^2+C_2^3+C_2^4+\cdots+C_2^9$,它们中的每个数已"在第二行"中计算出.这里的"行"指的是穆恩依姆把这些结果列成的一张表格中的行.表中第一行列出的数是 1,

وهكذا تخطيط المثال في الجدول											جدول جميع الجداول
من عشرة ألوان										1	1
جدول الشراريب التي من تسعة ألوان تسعة ألوان									1	9	10
جدول الشراريب التي من ثمانية ألوان ثمانية ألوان								1	8	36	45
جدول الشراريب التي من سبعة ألوان سبعة ألوان							1	7	28	84	120
جدول الشراريب التي من ستة ألوان ستة ألوان						1	6	21	56	126	210
من خمسة ألوان خمسة ألوان					1	5	15	35	70	126	252
من اربعة ألوان اربعة ألوان				1	4	10	20	35	56	84	210
من ثلاثة ألوان ثلاثة ألوان			1	3	6	10	15	21	28	36	120
من لونين لونين		1	2	3	4	5	6	7	8	9	45
من لون لون	1	1	1	1	1	1	1	1	1	1	10
	لون أول	لون ثان	لون ثالث	لون رابع	لون خامس	لون سادس	لون سابع	لون ثامن	لون تاسع	لون عاشر	جميع الألوان

图 7.13　伊本·穆恩依姆的组合表格.

(来源:A · Djebbar, Publications Mathematiques D'Orsay, 1985.)

2,…,10(= $C_1^1,C_1^2,\cdots,C_1^{10}$),第二行列出的是 1,3,6,…,36(= $C_2^2,C_2^3,C_2^4,\cdots,C_2^9$),等等.伊本·穆恩依姆不断地提到这张表即帕斯卡三角形,以说明各种计算可以如此容易地做出来(图 7.13).当他计算 $C_k^n,n\leqslant 10,k\leqslant n$ 时他逐行地说明了这张表格.那么,他指出

$$C_k^n = C_{k-1}^{k-1}+C_{k-1}^{k}+C_{k-1}^{k+1}+\cdots+C_{k-1}^{n-1},$$

至少对他的那些 n 和 k 如此.

回到文字的问题上,伊本·穆恩依姆也处理了排列问题:

> 问题是:我们需要定出一个典型的步骤来确定一个词的字母的排列数,其中字母的个数是给定的且没有重复.如果此词有两个字母,显然有两个排列,因为第一个字可以做成第二个而第二个做成第一个.如果我们再增添一个字母来考虑三个字母的词,清楚的是在两个字母的词中这两个字母的每个排列前面、中间和后面的位置均可放入第三个字母.因此三个字母的词中的字母有六个排列.如果现在在词中增添另一个字母构成一个四个字母的词,那么第四个字母可在这六个排列中的每个之中(四个位置中的一个).因此四个字母的词有二十四个排列.[30]

伊本·穆恩依姆得出结论说,不管一个词有多长,它的字母的排列数是 1 乘以 2 乘以 3 乘以 4 乘以 5 等等直至此词的字母数.伊本·穆恩依姆还寻求解决其他的许多问题,包括考虑到发音及母音符号而涉及具重复可能的排列.他的目的是确定可能的阿拉伯文字的个数,而且在讨论了这方面的确切意思之后,他用上述的某些想法明晰地算出了有九个字母的词的个数,其中每个词有两个不重复的字母,两个字母重复了两次,一个字母重复了三次.原来这是一个 16 位的数.

7.3.2　组合学和数论

13 世纪末,组合公式的问题为卡米勒·丁·法雷西(1320)所吸纳,他生活在波斯,这一次,问题却与整数分解因子和亲和数的概念有关.回想欧几里得在《原本》IX 中曾指出如何求完全数,即等于其真因子和的那种数.后来希腊的数学家推广了这个思想,定义了亲和数的概念,这是一个数偶,它的每一个等于另一个的真因子之和.可惜希腊人只发现了一个这样的数偶,即 220 和 284,没有能找到一个一般的定理来产生这样的数偶.是伊本·库拉首先发现并证明了这样一个定理,用现代的记号可表述如下:

伊本·库拉定理　对 $n>1$,令 $p_n=3\cdot 2^n-1,q_n=9\cdot 2^{n-1}-1$.如果 p_{n-1},p_n 和 q_n 为素数,则 $a=2^n p_{n-1}p_n$ 与 $b=2^n q_n$ 为亲和数.

作为最简单的例子,我们取 $n=2$,因为 $p_1=5,p_2=11,q_2=71$ 均为素数,得到的数偶为 220,284.虽然其他的伊斯兰数学家研究过伊本·库拉的结果,然而一直到了 13 世纪后期,第二个亲和数 17 296 和 18 416 才被法雷西找到,这与他自己对此定理的研究有关.

法雷西关于库拉定理的工作是通过组合分析进行的,在这里是指一个数的素因子的组合.正是这些组合确定了一个数的所有真因数.例如,如果 $n=p_1p_2p_3$,其中每个 p_i 都为素数,则 n 的因数为 $1,p_1,p_2,p_3,p_1p_2,p_1p_3,p_2p_3$,和 $p_1p_2p_3$.因此总共有 $C_0^3+C_1^3+C_2^3+C_3^3$ 个因数.因此,这种组合数间的关联知识对于整数因子的彻底研究是必不可少的.

法雷西用类似于伊本·穆恩依姆的推理较详细地做出了这些关联。事实上他也发展了“帕斯卡”三角形,不仅把这些行与组合数联系起来而且还与垛积数相联系,包括对三角形,棱锥体和更高阶立体的垛积,同时给出了库拉定理的一个代数证明.

7.3.3　伊本·班纳和组合公式

像他的先辈那样,法雷西以取和的方式来发展组合的结果.伊本·班纳(Abu-l'-Abbas Ahmad al-Marrakushi ibn al-Bannā,1256—1321),也是一位马拉喀什的数学家,可以说是伊本·穆恩伊姆在摩洛哥的直接后继人.他推导出了求组合数的标准乘法公式,此公式在印度已早有陈述.另外,他是以抽象方式来处理组合问题的,并不涉及哪一类具体对象被组合.

伊本·班纳一开始用了计数的推理证明了 $C_2^n = \frac{n(n-1)}{2}$:一个元素 a_1 与 $(n-1)$ 个元素中每个相伴,a_2 与 $n-2$ 个中每个相伴,等等,故 C_2^n 等于 $n-1, n-2, n-3, \cdots, 2, 1$ 的和.然后他指出要求 C_k^n 值"我们将先于所求组合值的前一个组合乘以一个这样的数,它在已给数之前并与它的距离等于所求组合的数目.由此乘积除以组合数目."[31] 我们可以把伊本·班纳的语句翻译为现代公式

$$C_k^n = \frac{n-(k-1)}{k} C_{k-1}^n.$$

在证明此结果时,伊本·班纳先从 C_3^n 开始.对这 n 个之中每个两元的集,配上剩下的 $n-2$ 个元中的每一个.这样便得到 $(n-2)C_2^n$ 个不同的集合.但因为 $C_2^3 = 3$,每个这样的集合被重复了三次.例如 $\{a,b,c\}$ 会以 $\{\{a,b\},c\}$, $\{\{a,c\},b\}$ 和 $\{\{b,c\},a\}$ 的样子出现.因此像断言的那样有 $C_3^n = \frac{n-2}{3}C_2^n$.下一步,我们知道 $C_3^4 = 4$.由此得出,如果我们把每个集合配上剩下的 $n-3$ 个元中的每一个,则总数 $(n-3)C_3^n$ 是 C_4^n 的四倍,或 $C_4^n = \frac{n-3}{4}C_3^n$.相似的推理对 k 的其他值也成立.把这些结果汇总起来,便得到

$$C_k^n = \frac{n(n-1)(n-2)\cdots(n-(k-1))}{1\cdot 2\cdot 3\cdot \cdots \cdot k},$$

这是 n 个元集合中取 k 个元的不同方式的个数的标准公式.用此结果以及伊本·穆恩依姆关于 n 个对象集合的排列数为 $n!$ 的结果,伊本·班纳用乘法证明了 n 元集合中取 k 个的排列数 P_k^n 为

$$P_k^n = n(n-1)(n-2)\cdots(n-(k-1)).$$

伊本·班纳对 C_k^n 公式的证明以及伊本·穆恩依姆对排列规律的证明像凯拉吉和塞毛艾勒的较早的证明那样,"具有归纳风格.就是说,作者由对一个小的值的已知结果开始,然后用它逐步地建立较高的值的结果.但是既非伊本·班纳也非他的任何一位先辈曾明晰地叙述过作为证明基础的归纳原理.第一次作这种叙述的是L·B·热尔松(Levi ben Gerson),他是伊本·班纳的一位同时代的年轻人,我们将在第8章说到他.

7.4　几何学

伊斯兰数学家一开始处理的是实用的几何,但以后就对此学科的许多理论方面作了研究,其中包括欧几里得的平行公设,无理量概念,确定立体体积的穷竭原理.

7.4.1　实用几何

像最早的代数那样,现存最早的阿拉伯几何书是花拉子米的著作,并且是作为其代数著作的单独章节出现的.比起他代数中的几何证明来说,在花拉子米的几何中简直看不到希腊理论数学的影

响.他的书是对诸如观测者所需的测量学规则的初级汇编.它没有公理也没有证明(除了对等边直角三角形的毕达哥拉斯定理的证明),它看起来非常像是古老的希伯莱几何《度量之书》,它在第4章中已讨论过.

我们先讨论花拉子米对圆的规则:

> 在任一圆中,它的直径与三又七分之一的积等于周长.这是由实际生活中得出的一般规则,但并不十分准确.几何学家们有两种其他的方法.其中之一是,先以直径自乘,然后乘以十;之后再对此乘积取根,此根便是周长.另一个是由一些天文学家使用的方法.方法是,你把直径乘以三万二千八百三十二,然后用二万去除此乘积.此商便为周长.两种方法几乎效果相同…….任一圆的面积由以直径的一半乘周长的一半得到,这是因为,在每个等边等角的多边形中……面积可以由通过它的中间那个圆的直径的一半乘周长的一半得到.如果你对任一圆的直径作自乘,并从此积中减去此积的七分之一与另一个七分之一的一半,则剩下的便是这圆的面积.[32]

这里的对 π 的第一个近似值是阿基米德的那个,即 $3\frac{1}{7}$,它类似于海伦(Heron)的和《度量之书》中的近似值.事实上,花拉子米给出计算圆面积的例子与《度量之书》中的一样,直径7的平方为49,然后减去它的七分之一和七分之一的一半,或者说是 $10\frac{1}{2}$,得到 $38\frac{1}{2}$.以 $\sqrt{10}$ 来近似 π 应归于几何学家,它在印度被使用.有意思的是,它较"不十分准确"的 $3\frac{1}{7}$ 更不准确.最早知道第三个近似值3.1416的也是印度人,出现在阿耶波多的著作中.把这个值归功于天文学家大概是联系到它在印度人写的天文学著作中曾被用过而此著作被翻译成了阿拉伯文.人们可以想像,花拉子米对后一个近似值曾请教过几何学家,这件事意味着,他知道对它们的某种证明.但是他并没有在他的几何著作中给出任何这类证明.

对菱形的面积公式简单地给出为:"你可以从两个[对角线]……来计算[菱形的]面积……你以一条对角线的一半去乘另一条对角线."[33] 这里给出来的例子与希伯莱著作中的一样,即一个边长为5而对角线长为6和8的菱形.面积从而为24.

对三棱台体积的讨论也与《度量之书》中的十分相似,而且例子又是一样的.它不是像莫斯科纸草书那样直接给出一个公式,花拉子米用相似三角形计算了到完整的三棱锥顶点的高,然后从下面三棱锥的体积中减去上面的那个三棱锥的体积.

但是在边长为13,14和15的三角形面积的计算上,花拉子米的做法不同于希伯莱著作.他不是使用海伦的公式,而是从相对的顶点向长度为14的边引一垂线,把垂足到底边一端的距离当作未知量 x,然后两次运用毕达哥拉斯定理来计算三角形的高(图7.14).于是,$13^2 - x^2 = 15^2 - (14 - x)^2$,$x = 5$,$h = 12$,从而面积为84.

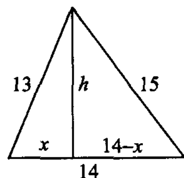


图7.14　花拉子米对一个三角形面积的计算.

7.4.2　平行公设

花拉子米的几何的确可被归类为"实用的".然而伊斯兰的作者们却受到他们阅读希腊理论著作的巨大影响,并很快由于他们所了解的欧几里得和其他作者的知识而对纯几何问题感到兴趣.其中在伊斯兰几何学中反复出现的一个思想是关于平行线和欧几里得第五公设的可证明性.甚至在

希腊时期,数学家们就为此公设而感到困扰.曾做过许多努力从其他公设来证明它.在伊斯兰世界里也是如此.例如,伊本·海塞姆在他的《对欧几里得原本中前提的评注》中,试图重新制定欧几里得的平行理论.他首先重新定义了平行线的概念.欧几里得定义平行直线为两条永不相交的直线,但伊本·海塞姆的"更明显的"定义包含了这种直线的可构造性假定.他写道,如果一条直线使其一端总在第二条直线上并使它垂直于第二条直线移动,则这条动直线的另一端就作出一条平行于第二条直线的直线.实际上,这个定义刻划出平行线是相互等距的直线这一特性,同时也把运动的概念引进了几何.后来的评论者们,其中包括奥玛·海亚姆,对此感到不满意.他们对一条移动并总与一条已知直线保持垂直的直线这种"不证自明"的性质表示怀疑,因而他们不能接受基于这种思想的证明.如他们所知,欧几里得只在从旧的对象生成新的对象时用过运动的思想,如像球面是由一个旋转的半圆生成的.但是,伊本·海塞姆却把这个思想用于他对第五公设的"证明"中.

伊本·海塞姆的证明的关键步骤是下面的引理:

引理 如果在一固定直线段的两个端点上作两条与此直线夹直角的直线,从这两条直线之一上任一点作垂直于另一条直线的直线段,则每一条这样的线段都等于所给的固定直线段.

在图7.15中,GA与DB被画成与AB夹直角,并从G向直线DB引出一垂线.必须证明GD等于AB.伊本·海塞姆以矛盾法来证明.他先假设$GD > AB$.他延长GA过A使得$AE = AG$.类似地,延长BD过B.由点E向延长了的BD作垂线,交它于T.作直线GB与BE.三角形EAB与三角形GAB由边-角-边关系全等.从而$\angle GBA = \angle EBA$,故$\angle GBD = \angle EBT$.由此得出三角形EBT与GBD全等,因而$GD = ET$.现在运用他的移动观念,伊本·海塞姆想像直线ET沿直线TD移动,并保持总与它垂直.当T与B重合时,由于$ET > AB$,点E将在直线AB之外.我们称在这个特定时间上的ET等于HB.当然,当ET到达GD时,这两段直线重合.现在由平行性的定义得出直线GHE是一条平行于DBT的直线.由构造知道GAE也是一条直线,故而就会有具有相同端点的两条不同直线,从而此两直线包围了一块空间.当然这是不可能的.由假定$GD < AB$,也类似地引出矛盾.证完.

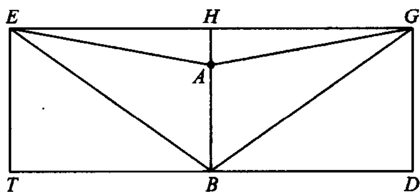


图7.15 伊本·海塞姆处理平行公设时的引理的证明.

因为$GD = AB$,容易得到$\angle AGD$也像四边形$ABDG$其他三个角一样也是个直角.于是,便不难证明欧几里得的公设.当然,伊本·海塞姆没有意识到的是他最初的平行性定义已经隐含了那个公设.不管怎样,他的结果澄清了平行公设与任意四边形角之和等于四个直角这个事实之间的等价关系.

奥玛·海亚姆也对平行性问题感兴趣.在他的《关于欧几里得书中有疑问的公设的评注》中,他以如下原理作为出发点,即两条收敛的直线相交,而它们在收敛的方向上不可能发散.他说收敛直线的意思是指相互趋近的直线.给出了此公设后,奥玛·海亚姆对归结于欧几里得第五公设的八个命题进行了证明.他先构造了一个具两条等长垂线AC与BD的四边形,这两条垂线是从一已知线段AB的两端作出的,然后连接C和D点(图7.16).他证明在C和D处两个角都为直角的办法是指出其他两种可能性,即它们均为锐角或均为钝角都将引起矛盾.如果它们为锐角,则CD就会长于AB,如果它们为钝角,则CD就会短于AB.对每个情形他都指出直线AC和BD会在

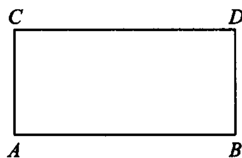


图7.16 奥玛·海亚姆的四边形:$AC = BD$, $AC \perp AB$,且$BD \perp AB$. C和D处的角是锐角,钝角还是直角?

AB 的两边同时发散或收敛,这与他的原始公设相矛盾.奥玛·海亚姆现在便可以证明欧几里得的第五公设了.在某种意义上说,奥玛·海亚姆的处理较海塞姆的好,这是因为他明确地定出了一个新的公设来取代欧几里得的平行公设,而不像后者那样把它隐藏在一个新定义之中.

人物小传	纳西尔·丁·图西(1201—1274)(Nasīr al-Dīn al-Tūsī)
	纳西尔·丁来自伊朗的图斯,在波斯的尼萨普尔完成了他的正式教育.这个地方是当时的主要学术中心.不久作为一个学者,他赢得了很大的声誉.但是,13 世纪在伊斯兰历史上是个极为动荡的时期.伊朗惟一和平的地方是那些由伊斯玛仪派统治者们直接控制的城堡.所幸的是,纳西尔·丁说服了其中一位统治者让他在这样一个城堡里工作.在 1256 年蒙古的头领忽拉古打败了伊斯玛仪派之后,图西转变了他的效忠对象.他作为一位科学顾问服务于忽拉古并获得了他的赞助,在马拉格哈建造了一座观测站.它建在大不里士南方 50 公里的一个城镇.在这里纳西尔·丁作为一个庞大的天文集体的领头人度过了他的余生.这段时间里,他计算了一组非常精确的天文表格并发展出一种天文模型,这个模型可能曾被哥白尼采用来设计他的日心体系.



图 7.17　纳西尔·丁·图西.

在奥玛·海亚姆之后大约一个世纪,另一位数学家纳西尔·丁·图西(Nasīr al-Dīn al-Tūsī,1201—1274)对他的先辈们的工作作了详细的批评,并且在他 1250 年左右的书《抹去平行线方面疑点的讨论》中也努力做出了自己对第五公设的证明.他考虑了与奥玛·海亚姆一样的四边形,也试图从锐角和钝角的假定中导出矛盾.但是在一篇论文中他有一个基于另一假定之上的新推理,这个假定等价于欧几里得的公设.这篇长文大概是他的儿子萨德尔·丁根据纳西尔·丁对这个主题的后期思想于 1298 年写成的.这个假定说,如果一条直线 *GH* 在 *H* 垂直于 *CD*,并在 *G* 斜交于 *AB*,则从 *AB* 向 *CD* 引的垂线如果在 *GH* 与 *AB* 交出钝角的那一侧,则大于 *GH*,而如果在另一侧,则小于 *GH*(图 7.18).这后一个工作的重要性在于它于 1594 年在罗马被发表,并被欧洲的几何学家们所研读.特别地它成了萨凯里(Saccheri)工作的起点并最终导致了非欧几何的发现.[34]

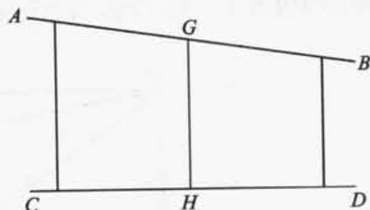


图 7.18　纳西尔·丁关于平行线和垂线的假设.

7.4.3　不可公度性

使伊斯兰数学家感兴趣的另一个几何主题是不可公度性.关于欧几里得《原本》X 的专题有许多阿拉伯文的评论.回想一下,伊斯兰的代数学家早就开始在他们关于方程式的工作中使用了无理量,他们在那里忽略了欧几里得在数和量之间的区别.但是,有许多评论家努力把这个用法放进了一个理论框架之中,这个框架与欧几里得的工作是保持一致的.

大约在 1000 年左右,伊本·巴格达蒂(Abū ʻAbdallāh al-Hasan ibn al-Baghdādī)写了一本书《论可公度和不可公度》,他试图把对无理量使用过的运算规则和《原本》的主要原理协调一致,因而证明同代人的计算方法是有效的.他完全明白,这些数值的计算方法较之于欧几里得的几何模式更简单:"采用一个数并将自己置于它的基础上较之于对一个量作同样的假定 …… 要更容易一些."[35]因为他了解亚里士多德和欧几里得对数和量之间所作的基本区别,所以他一开始便用建立数与线

段之间的对应关系的办法把这两种概念联系起来,这个方法看起来像是现代的方式.这就是说,给出一个单位长度 a,每个"整数"n 对应于单位长度的适当倍 na.这个长度的部分量,如像$(m/n)a$,则对应于一个数的部分(m/n).伊本·巴格达蒂把任一依此种方式表达的量考虑为一个有理量.他指出如《原本》X-5中所说,这些量间的相互关联就像数与数之间一样,那些不是"部分"的量可以考虑为无理量.实际上,伊本·巴格达蒂试图把有理数嵌入到一条数直线中.但他也想要把无理量与"数"联系起来.

伊本·巴格达蒂通过根式的想法来做这个联系.一个数 n 的根是连续比例$n:x=x:1$的中项.这样的有理根可能存在也可能不存在.然后他按相似的方法定义量 na 的根为单位长度a与长度na的比例中项.这个量总能用直尺和圆规构造出来,故而它必定存在.自然它或是有理的或是无理的.由于"有理数"对应于"有理量"而且由于后者总有根,这根或是有理的或不是,故而他可以考虑前者的根来延续这个对应关系.特别地,他注意到对长度而言,根与平方具有相同的几何类型.换句话说,表示为线段的量的根是另一线段,这正与一条线段的平方可表示为一条线段一样.伊本·巴格达蒂如同他的一些伊斯兰前辈,离开了希腊人坚持的齐性而走向这样一个观念,即所有的"量"可以用相同的方式在本质上表达为"数".

伊本·巴格达蒂以广泛论述欧几里得在《原本》X中处理过的各种无理量作为全书的结尾.这些讨论的一个结果是,他证明了无理量的"稠密性",即在任意两个有理量之间存在无穷多个无理量.例如,他考虑了由相邻数2和3所表示的量.这些量的平方由4和9代表.这两个量之间为由数5,6,7和8所代表的量.它们的根$\sqrt{5}$,$\sqrt{6}$,$\sqrt{7}$和$\sqrt{8}$被巴格达蒂称为2与3之间的第一阶无理量.类似地,4和9的平方即16和81,以及25,36,49和64也都代表了量.对应于整数17,18,…,24的第一阶无理量为$\sqrt{17},\sqrt{18},\cdots,\sqrt{24}$,同样第二阶无理量为$\sqrt{\sqrt{17}},\sqrt{\sqrt{18}},\cdots,\sqrt{\sqrt{24}}$.后面的这些量在原来的量2与3之间.伊本·巴格达蒂指出,人们可以按此方式继续下去找到在原来给出两个数之间的任意多个具各种阶的无理量.伊本·巴格达蒂的著作表明了伊斯兰的作者们既理解了他们希腊前辈们坚持把量和数这两个领域分割开但也要求打破这种一分为二的束缚,从而使他们在计算中大量使用"无理数"成为合理.

7.4.4 体积和穷竭法

我们就要讨论的这最后一个几何领域也证明了伊斯兰的作者们了解希腊人的著作并想要超越他们.这是指他们用由欧多克索斯所开创而被阿基米德所广泛应用的穷竭法来计算立体体积的工作.原来,虽说伊斯兰数学家读过阿基米德的著作《论球面和圆柱》,他们却没有机会读到《论圆锥体和球体》,在此书中阿基米德指出如何计算一条抛物线绕轴旋转形成的立体的体积.因而塔比·伊本·库拉找出了他自己的证明,它十分冗长而复杂;大约在75年之后,一个来自里海南部地区的名叫阿布·萨赫勒·库西(Abū Sahl al-Kūhī,10世纪)的人简化了塔比的方法,并解决了一些关于体积的问题以及一些关于重心的类似问题.不久之后,便轮到库西挨批评了.伊本·海塞姆批评他没有在所有的一般情形下解决抛物体的问题,这指的是没有考虑到一段抛物线绕垂直于它的轴的直线旋转成的立体的体积.正是这后一个问题由海塞姆本人进行了解答.

用现代术语说,伊本·海塞姆证明了把抛物线 $x=ky^2$ 绕直线 $x=kb^2$(它垂直于抛物线的轴)旋转得到的立体的体积等于半径为 kb^2,高为 b 的圆柱体积的8/15.他的形式推理是个典型的穷竭推理.就是说,他假定所求的体积大于那个圆柱体积的8/15然后推导出矛盾;再假定它小于那个8/15的体积并又推导出矛盾,但是伊本·海塞姆推理的实质涉及把圆柱"片"成 n 个厚度为$h=b/n$

的圆片，每个圆片与抛物体的相交部分给出了抛物体一片的体积的近似值（图7.19）．在抛物体中的第 i 个圆片具半径长 $kb^2 - k(ih)^2$，从而具体积 $\pi h(kh^2n^2 - ki^2h^2)^2 = \pi k^2h^5(n^2 - i^2)^2$．抛物体的总体积因此近似地等于

$$\pi k^2h^5\sum_{i=1}^{n-1}(n^2 - i^2)^2 = \pi k^2h^5\sum_{i=1}^{n-1}(n^4 - 2n^2i^2 + i^4).$$

但是伊本·海塞姆已经知道了整数平方和与整数四次幂和的公式．用这些公式，他能够算出

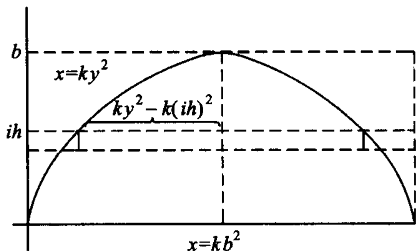


图7.19　绕垂直于其轴旋转的一段抛物线．

$$\sum_{i=1}^{n-1}(n^4 - 2n^2i^2 + i^4) = \frac{8}{15}(n-1)n^4 + \frac{1}{30}n^4 - \frac{1}{30}n$$
$$= \frac{8}{15}n \cdot n^4 - \frac{1}{2}n^4 - \frac{1}{30}n,$$

因此有

$$\frac{8}{15}(n-1)n^4 < \sum_{i=1}^{n-1}(n^2 - i^2)^2 < \frac{8}{15}n \cdot n^4.$$

但是外围圆柱的一个标准片的体积是 $\pi h(kb^2)^2 = \pi k^2h^5n^4$，因此圆柱的总体积为 $\pi k^2h^5n \cdot n^4$，而除去它的"顶片"的那个圆柱体积为 $\pi k^2h^5(n-1)n^4$．故而，这个不等式表明，抛物体的体积界于除去顶片的那个圆柱体积的8/15与整个圆柱体积的8/15之间．因为顶片可以随 n 充分大而随意地小，这更推导出所断言的结果，即抛物体的体积恰是那个圆柱体积的8/15．

7.5　三角学

在8世纪后期，一本印度的《悉檀多》（"历数书"）被带到了巴格达并翻译成了阿拉伯文．这使得伊斯兰学者得以了解到印度人的三角学知识，这些知识早先是从希帕科斯的希腊文本翻译改编得到的．当托勒密的《大成》被翻成阿拉伯文后不久，伊斯兰学者们便注意到了在那书中详细论述的三角学．像在其他数学领域中那样，伊斯兰数学家吸取了他们从其他文化中发现的东西，并逐步地在此学科中灌输进他们自己的思想．

像在希腊和印度的情形一样，在伊斯兰三角学也是紧密地与天文学联结在一起的，故而一般情形下，关于三角学的数学教科书总被写成更为广泛的天文学著作中的章节．数学家特别有兴趣用三角学去解球面三角形，因为伊斯兰法律要求穆斯林在祈祷时必须面向麦加的方向．在一个人自己的位置上要确定适当的方向需要解决在地球球面上这种三角形解的广泛知识．平面和球面三角形的解对于确定祷告者的正确时刻也是重要的．一般地，这些时刻的确定关系到黎明的第一道阳光和黄昏的最后一缕阳光，还有白天的长度，以及在指定日子里太阳的高度，而这些概念也需要借助球面三角形来精确地决定．

7.5.1　三角函数

回想一下，托勒密在他的三角学著作中只用了一个三角"函数"即弦，而印度人把它改进为更加方便的正弦(sine)．在伊斯兰三角学初期，弦与正弦是同时使用的，而最终还是正弦胜出．（像印度人那样，弧的正弦在伊斯兰是在一个半径为 R 的圆中一条特定的直线．）还不完全清楚谁首先引进了其他的函数，但我们的确知道巴塔尼（Abū ʻAbdallāh Muhammad ibn Jābir al-Battānī，855—929）

在他为改进《天文学大成》而写的天文学著作中使用了“对 90° 的余角的正弦”(即我们所称的余弦).因为他不用负数,故他只对小于 90° 的弧定义了余弦.对介于 90° 到 180° 的弧他使用了正矢函数 versine,定义为 $\text{versin}\ \alpha = R + R\sin(\alpha - 90°)$.由于巴塔尼没有使用正切,故而他的公式与托勒密的同样粗糙.

人物小传

比鲁尼(973—1055)(Al-Bīrūnī)

比鲁尼(图 7.20),生于花拉子模的一个靠近现在乌兹别克斯坦叫作比鲁尼的市镇的地方.很小的时候,就在阿布·纳斯尔·曼舒尔·伊本·伊拉克的指导下开始学习科学.伊拉克是这个地区中一位卓越的天文学家.他家乡的政治倾轧迫使他于 995 年逃离,但两年之后又回到了花拉子模的主要城市喀瑟,去观察一次日蚀.他预先安排了阿布·瓦法(Abu'l-Wafā')在巴格达观察同一次日蚀,故而这两处发生的事的时间差使他能够算出这两地的经度差.1017 年,阿富汗加慈纳的苏丹马赫穆德征服了花拉子模,并很快统治了包括北印度在内的一个广大的王国.比鲁尼被带到了苏丹的宫庭,然后又旅行去了印度,在那里他写了一本关于印度文化各个方面的主要著作,它包括了各个不同的专题,如像种姓制度,印度教的哲学,象棋规则,时间的观念以及历法的程序.比鲁尼总共写了超过 140 部著作,其中大部分是关于数学,天文和地理的.



图 7.20 叙利亚邮票上的比鲁尼.

正切,余切,正割和余割函数出现在伊斯兰著作中是在 9 世纪,或许最早是在哈西卜(al-Hāsib,770—870)的著作中,虽说在 8 世纪时正切函数已在中国被使用了.然而,在这里我们所要考虑的是比鲁尼在他的《测影通论》(Exhaustive Treatise on Shadows)中对这些函数的讨论.“正影子[余切]的一个例子是:设 A 为太阳,BG 为垂直于 EG 的日规指针,EG 平行水平面,而 ABE 为通过指针顶头的日光光线(图 7.21(a))……,EG 就是那个被称作正影子的线段,它的基点为 G 而尾端为 E.连接影子和指针顶端的直线 EB 是这影子的斜边[余割].”[36]

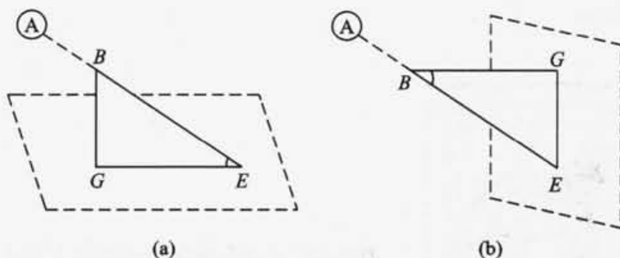


图 7.21 比鲁尼对正切、余切、正割、余割的定义.(a) 中 GE 是角 E 的余切,而 EB 是余割;(b) 中 GE 是角 B 的正切而 BE 是正割.

相似地用平行于水平面的日规指针定义正切和正割.在图 7.21(b) 中,GE 被称作“逆影”(正切)而 BE 被称作“逆影的斜边”(正割).

比鲁尼证明了三角函数之间的各种关系.例如,他指出“指针与影子斜边之比同于高的正弦与总正弦之比.”[37] 比鲁尼说的“总正弦”是指 90° 弧的正弦即取弧的那个圆的半径.那么此公式可以翻译成

$$\frac{g}{g\ \csc\alpha} = \frac{R\sin\alpha}{R}$$

(这里 g 表示指针的长度),或者表示为

$$\csc\alpha = \frac{1}{\sin\alpha}.$$

比鲁尼进一步指出"如果我们已知在某时刻的影子，而我们想要求在那个时刻的太阳的高度，我们将影长乘以影长的等量又将指针长乘以它的等量并取和[的平方根]，它就是余割.然后以它去除指针长和总正弦的乘积，这便得到高度的正弦.我们再由正弦表查出它对应的弧，就得到了在影子的那个时刻的太阳高度."[38] 用现代的记号，表明比鲁尼运用了关系式

$$\sqrt{g^2\cot^2\alpha + g^2} = g\csc\alpha \quad (\text{或 } \cot^2\alpha + 1 = \csc^2\alpha)$$

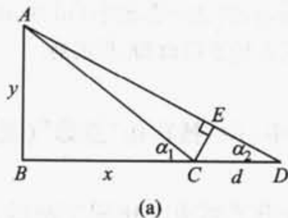
和前面的那个公式$\frac{gR}{g\csc\alpha} = R\sin\alpha$来确定以所用的特定半径值$R$为基础的正弦函数，然后参照他的正弦表从逆向定出 α.比鲁尼类似地给出了等价于 $\tan^2\alpha + 1 = \mathrm{scc}^2\alpha$ 和 $\tan\alpha = \frac{\sin\alpha}{\cos\alpha}$ 的规则并且给出了正切和余切的一张表，对此表他使用了关系 $\cot\alpha = \tan(90° - \alpha)$.

收集在他书中的大量的三角学知识只被比鲁尼用于处理天文学问题，这或许有些令人惊奇.对于确定地面上的高度和距离，他所描述的是非三角学的方法.例如确定清真寺旁光塔的高度，而且是塔基可以接近的情形，他启发说"如果在太阳的高度等于一次旋转的八分之一[45°]的时刻进行观察，则影子的末端和垂足之间的距离等于[它的高度]."[39] 如果塔基是不能接近的，比鲁尼则描述了一个类似于第 6 章中讨论过的中国和印度的方法.但是与他的中国和印度前辈不一样，他在书中给出了用相似三角形思想的推理.

但是在《测影通论》出现前的四分之三世纪时，卡比西(Abu l-Saqr al-Qabisi，10 世纪)已经描述过一种只用正弦来确定不可及对象的高度和距离的三角方法.人们可用一个星盘(通常用于天文观测的一种量角工具)在两个地方 C、D 观测顶点 A，角 $\alpha_1 = \angle ACB$，$\alpha_2 = \angle ADB$(图 7.22a 和 b).如果 $CD = d$，则高 $y = AB$，距离 $x = BC$ 由下面公式给出：

$$y = \frac{d\sin\alpha_2}{\sin(90° - \alpha_2) - \frac{\sin(90° - \alpha_1)\sin\alpha_2}{\sin\alpha_1}},$$

$$x = \frac{y\sin(90° - \alpha_1)}{\sin\alpha_1}.$$





(b)

图 7.22　(a) 由确定两个角的方式来确定高度和距离的卡比西方法.(b) 伊朗邮票上的星盘.

7.5.2　球面三角学

虽然有几个把三角学用于地面测量的例子，但是三角函数的主要应用还是在于解由天文学问题而来的球面三角形.伊斯兰数学家们比起托勒密处理这些问题来能够推导出更为简便的方法.基

本的结果似乎是由两个与比鲁尼同时代的人各自独立发现的,一个是比鲁尼的一位老师,名叫伊本·伊拉克(Abū Nasr Mansūr ibn 'Iraq,1030),另一个是阿布·瓦法(Muhammad Abū'l-Wafā' al-Būzjānī,940—997),他是巴格达的一位重要天文学家.我们将按照后者的天文学手册《天文学大全》来谈谈他的工作.

第一个是被后来称为"四量规则"的结果.

定理　如果 ABC 和 ADE 是两个球面三角形,各自在 B 和 D 为直角,并在 A 有一公共锐角,则 $\sin BC : \sin CA = \sin DE : \sin EA$(图 7.23).

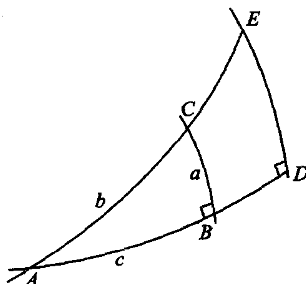


图 7.23　"四量规则".

这个定理的一个直接推论是第 4 章中讨论过的梅纳劳斯(Menelaus)定理的一个特殊情形:如果 ABC 是个 B 为直角的球面直角三角形,则 $\sin A = \frac{\sin a}{\sin b}$.为证明这点,延长斜边 AC 和底边 AB 各至点 E 和 D,使得 AD 和 AE 都是大圆的四分之一.从 E 到 D 的大圆的弧垂直于 AD 和 AE 两者,于是可应用定理.因为 $\sin DE = \sin A$,故而我们的结果得证.这个推论实质上被托勒密在他的许多计算中用过.阿布·瓦法也给出了梅纳劳斯定理的其他特殊情形,包括了如像 $\frac{\cos a}{\cos b} = \frac{1}{\cos c}$ 和 $\frac{\sin c}{\tan a} = \frac{1}{\tan A}$ 的结果.另外,他给出了任一球面三角形正弦定理的证明.

定理　在任一球面三角形 ABC 中,$\frac{\sin a}{\sin A} = \frac{\sin b}{\sin B} = \frac{\sin c}{\sin C}$(图 7.24).

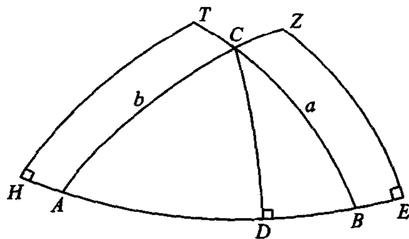


图 7.24　阿布·瓦法对正弦定理的证明.

给出球面三角形 ABC,令 CD 为垂直于 AB 的大圆弧.延长 AB 和 AC 至 AE 和 AZ,使其都为四分之一大圆,延长 BA 和 BC 至 BH 和 BT,也使它们为四分之一大圆.那么 A 是对大圆 EZ 的一个极点,B 是对大圆 TH 的一个极点.因为在 E 和 H 的角为直角,从而得到三角形 ADC 与 AEZ 为具有公共角 A 的球面直角三角形,而三角形 BDC 与 BHT 为具公共角 B 的球面直角三角形.由四量规则,我们有

$$\frac{\sin DC}{\sin b} = \frac{\sin ZE}{\sin ZA} \quad 及 \quad \frac{\sin DC}{\sin a} = \frac{\sin TH}{\sin TB}.$$

但因为 A 和 B 各为 ZE 和 TH 的极点,弧 ZE 等于 $\angle A$ 而且弧 TH 等 $\angle B$.因此这些等式可重写为形式:

$$\frac{\sin DC}{\sin b} = \frac{\sin A}{R} \quad 和 \quad \frac{\sin DC}{\sin a} = \frac{\sin B}{R}.$$

因此 $\sin A \sin b = \sin B \sin a$,从而正弦定理得证.

有了正弦定理之后,比鲁尼便能证明如何确定"奇伯拉"(qibla),即一个人所处地方的朝麦加的方向.对此问题,比鲁尼的一个解可描述如下.[40] 假定 M 是麦加的位置,P 是某人当前所在的地方(图 7.25).设弧 AB 表示赤道,T 为北极,并经 T 分别作通过 P 和 M 的经线.于是"奇伯拉"便是在

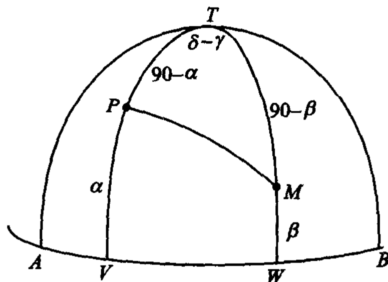


图 7.25　"奇伯拉"问题.

地球表面上的 $\angle TPM$. 假设已知 P 和 M 的纬度各为 α, β, 经度各为 γ, σ, 于是便知道了弧 TP 与 TM(各为 $90° - \alpha, 90° - \beta$), 也知道了 $\angle PTM (= \sigma - \gamma)$. 可惜正弦定理本身并不足以解出三角形 PTM, 这是因为不知道它的任一个角及其对边. 但是比鲁尼却对一系列的三角形反复地用此定理来达到目的.

我们按比鲁尼的方法, 作为例子取 P 为耶路撒冷(纬度 31°47′N, 经度 35°13′E). 麦加的纬度为 21°25′N, 经度 39°49′E. 让圆 $KSQN$ 表示从上方看时 P 的水平圆(或它的局部天顶), M 表示麦加的天顶(图 7.26). 如果 S 是水平面的南向点(P 是 M 的西北向点), N 是北向点, 作弧 PMK 和 NPS, 而弧 NK 则代表了"奇伯拉". 设圆 CFD 代表麦加的水平圆, 圆 MHJ 代表了 F 的水平圆, 并且通过北方天极 T 作圆 MTL. 问题的这些数据给出了 $TN = \alpha = 31°47'$, $TL = \beta = 21°25'$, $MT = 90° - \beta = 68°35'$, 和 $\angle MTH = \sigma - \gamma = 4°36'$. 由于已知 MT, $\angle MTH$ 及 $\angle THM = 90°$, 则对三角形 MTH 的正弦定理表明

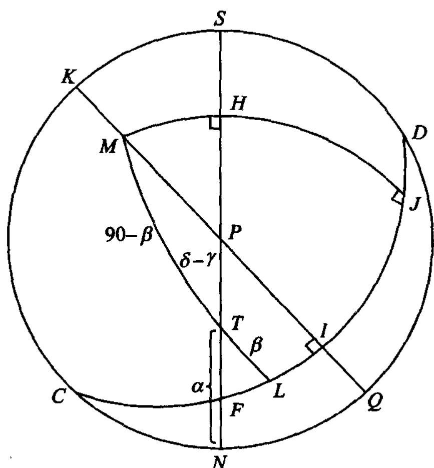


图 7.26　比鲁尼对"奇伯拉"的解.

$$\sin MH = \frac{\sin MT \sin\angle MTH}{\sin\angle THM} = 0.074\,66 .$$

因此 $MH = 4°17'$, $HJ = 90° - MH = 85°43'$. 因为 $\angle TFL = HL$, TL 和 $\angle TLF = 90°$ 为已知, 那么正弦定理应用到三角形 TFL 便确定了

$$\sin TF = \frac{\sin TL \sin\angle TLF}{\sin\angle TFL} = 0.366\,17,$$

故而 $TF = 21°29'$, 因此 $FN = \alpha - TF = 10°18'$ 及 $PF = 90° - FN = 79°42'$. 下一步, 将四量规则用于三角形 FPI 和 FHJ. 又因为 PF, $FH = 90°$ 且 HJ 已知, 所以可确定 $\sin PI$ 为

$$\sin PI = \frac{\sin PF \sin HJ}{\sin FH} = 0.981\,14,$$

于是 $PI = 78°51'$, $IQ = 90° - PI = 11°9'$. 但 C 是圆 $KMPIQ$ 的极点. 因而知道了 $\angle FCN (= IQ)$. 最后, 对三角形 CFN 应用正弦定理. 因以下三个量又为已知, 即 $\angle FCN$, $\angle CFN (= \angle TFL)$, 及 FN, 故第四个量 NC 便被确定:

$$\sin NC = \frac{\sin\angle CFN \sin FN}{\sin\angle FCN} = 0.922\,04,$$

$NC = 67°14'$, 且"奇伯拉" $NK = NC + CK = 67°14' + 90° = 157°14'$.

正弦定理和四量定理及它们的各种推论, 还有各种函数的表, 使得伊斯兰数学家能够解决对天文学及相关宗教目的具有重要意义的球面三角形问题. 这些结果出现在伊斯兰世界各地的天文学著作中. 在 12 世纪, 这些球面三角学的基本定理, 甚至可在西班牙的贾比尔(Abū Muhammad Jābir ibn Aflah al-Ishbīlī, 12 世纪初)的书中找到. 除了知道他来自塞维利亚以外对贾比尔的生活几乎一无所知, 但是他的主要著作, 一本评论托勒密《天文学大成》的书在 12 世纪后期被翻译成了拉丁文, 它给欧洲人提供了关于托勒密三角学在伊斯兰的进展方面一个最早的描述.

在伊斯兰世界自身, 直到 13 世纪才出现了与天文学独立的一本关于球面和平面三角学方面系统而内容广泛的著作, 即《论横截图形》(通常它以《论完全四边形》而知名), 它是纳西尔·丁·图西写的. 在这本书中, 作者证明了平面三角形的正弦定理, 然后系统地用它去解平面三角形. 但是他并

没有提到在"模糊"情形即已知两边及一边的对角的情形下,有两个解的可能性.也没有提及余弦法则,而在此种情形我们总是使用这个法则,而他是将三角形分成两个直角三角形,并用标准的直角三角形方法来完成他的解.

纳西尔·丁的著作也包含了解球面三角形的规则.他不仅包括进了我们在第4章中提到的梅内劳斯的球面直角三角形的所有四种特殊情形,而且还加进了另外两个情形:$\cos c = \cot A \cot B$ 和 $\cos A = \cos a \sin B$.他指出了如何利用这六个结果来系统地解决所有类型的球面三角形.特别地,他指出在已知三条边的情形下该如何解,也第一次讨论了在已知三个角情形下的解.

7.5.3 三角函数表

在处理天文学和地理学的问题时,不仅需要解所需三角形的公式而且需要高精度的表.这些表是逐步发展起来的.例如比鲁尼计算了一个正弦表,其间隔为 15′ 而精确度为六十进制的四位.像托勒密的计算那样,表的精度主要依赖于 sin 1° 的计算精度.对此计算有各式各样的方法,令人印象最深的是 15 世纪初卡西的方法.他从三倍角公式 $\sin 3\theta = 3\sin\theta - 4\sin^3\theta$ 出发.令 $\theta = 1°$ 便得出 $x = \sin 1°$ 的一个三次方程,即 $3x - 4x^3 = \sin 3°$.因为卡西是基于半径 60 的圆来计算它的正弦表,故而他需要计算 $y = 60\sin 1° = 60x$.他的方程因而是 $3y - \frac{4y^3}{60^2} = 60\sin 3°$,或者是

$$y = \frac{900(60\sin 3°) + y^3}{45 \cdot 60}.$$

回想一下,运用差角和半角公式可以计算 sin 3° 到任意精度.事实上卡西使用的 60 sin 3° 的 60 进位的值是 3;8,24,33,59,34,28,15.写成 60 进位的记号,他的方程因此为

$$y = \frac{47,6;8,29,53,37,3,45 + y^3}{45,0},$$

他所知道的是这个值与 1 靠近,由此他用迭代步骤来解此方程.将此方程以符号表示为 $y = \frac{q + y^3}{p}$,并假定解由 $y = a + b + c + \cdots$ 给出,其中各个字母表示依次的 60 进制数值,并由第一个近似值 $y_1 = q/p \approx a(= 1)$ 开始.为求第二个近似值 $y_2 = a + b$,令

$$y_2 = \frac{q + y_1^3}{p} \quad 或 \quad a + b = \frac{q + a^3}{p}$$

来解 b.于是

$$b \approx \frac{q - ap + a^3}{p} \ (= 2).$$

类似地,如果 $y_3 = a + b + c$,令

$$y^3 = \frac{q + y_2^3}{p} \quad 或 \quad a + b + c = \frac{q + (a + b)^3}{p},$$

求出

$$c \approx \frac{q - (a + b)p + (a + b)^3}{p} \quad (= 49).$$

卡西没有说明这个逼近步骤的正确性,但他显然知道它较他的先辈用来解三次方程的步骤要收敛得快得多.这时他计算出 $y = 1;2,49,43,11,14,44,16,26,17$,等价于 10 进位制下的 sin 1° 的值 0.017 452 406 437 283 571,在电动计算器出现之前的时代这是个相当了不起的业绩.卡西的庇护

人乌鲁伯格(图 7.27),他本身也是一位天文学家,统治着以撒马尔罕为首都的中亚的一片区域.他利用这个工作计算了正弦和正切表,对每一分弧算到 60 进制的第五位,在每张表中总共有 5400 个分项.

到了卡西的时代,伊斯兰的科学已处于衰退状态,在未来的年代里几乎再没有其他重要的科学家出现了.然而甚至在 15 世纪之前,数学的活动已在欧洲得以复苏.这种复苏的核心因素是 12 世纪的翻译家们的工作,他们使欧洲人能得到伊斯兰数学文库中的一部分资料.这工作的其余部分可能已找到了通向欧洲的道路,而这条道路还未被充分理解,但可惜的是我们在本章中讨论的数学家们的一些先进的工作并没有及时地到达欧洲,使它们能在欧洲人进行他们自己的数学开发中得到利用.我们将在第 8 章和以后的章节中指出目前已经知道的那些从伊斯兰到欧洲的数学思想的传播.

图 7.27 土耳其邮票上的乌鲁伯格.

习 题

十进制算术方面的问题

1. 运用乌格里狄西的方法做 8023 乘以 4638.
2. 运用乌格里狄西的方法对 135 增加自身十分之一连续五次,并验证其结果值正如正文中给出的 217.418 85.

代数方面的问题

3. 花拉子米对他的第六种情形 $bx + c = x^2$,给出下列规则:取根的个数的一半.对这个数自乘.把这个平方加到方程的数上.对它开平方根.把这个加到根个数的一半上.那就是解.把这个规则翻译成公式.利用图 7.28 给出它的正确性的几何推理,其中 $x = AB = BD$, $b = HC$, c 由矩形 $ABRH$ 表示, G 是 HC 的中点,且 $HK = HG$.

图 7.28 花拉子米对 $bx + c = x^2$ 的解规则的正确性说明.

4. 应用适当的公式解下面花拉子米提出的问题:

   (a) $\left(\frac{1}{3}x + 1\right)\left(\frac{1}{4}x + 1\right) = 20$.

   (b) $x^2 + (10 - x)^2 = 58$.

   (c) $x/3 \cdot x/4 = x + 24$.

5. 对方程 $\frac{1}{2}x^2 + 5x = 28$ 乘以 2 然后用花拉子米的步骤解此方程.类似地,先除以 2 再解方程 $2x^2 + 10x = 48$.
6. 解花拉子米的问题:我把 10 分成两份,并用第二份去除第一份,并用第一份去除第二份,这些商的和是 $2\frac{1}{6}$.求各个部分.
7. 解下列阿布·卡米勒的问题:

   (a) 假设 10 被分为两部分,一部分的自乘等于另一部分与 10 的平方根的积.求各部分.

   (b) 假设 10 被分为两部分,它们的每一部分被另一部分除,这两个商之和等于 5 的平方根.求各部分.(阿布·卡米勒用两种方法解此问题,其一是直接求 x,而第二种是先设 $y = \frac{10 - x}{x}$.)
8. 解下面的阿布·卡米勒的问题:

(a)$[x-(2\sqrt{x}+10)^2]^2=8x$　(先作替换 $x=y^2$.)

(b)$(x+\sqrt{\frac{1}{2}x})^2=4x$　(阿布·卡米勒给出了三种不同的方法:第一,直接解 x,其次是作替换 $x=y^2$,最后是作替换 $x=2y^2$.)

(c)$(x+7)\sqrt{3x}=10x$　(阿布·卡米勒给出了两个解.)

9. 解阿布·卡米勒的下面涉及三个变量的问题:$x<y<z$,$x^2+y^2=z^2$,$xz=y^2$,$xy=10$(先设 $y=10/x$,$z=100/x^3$,并替换进第一个方程.)

10. 完成塞毛艾勒的 $20x^2+30x$ 除以 $6x^2+12$ 的步骤以得到正文中所叙述的结果.证明此商的系数满足规律 $a_{n+2}=-2a_n$,其中 a_n 是 $1/x^n$ 的系数.

11. 用塞毛艾勒的步骤把 $20x^6+2x^5+58x^4+75x^3+125x^2+96x+94+140\frac{1}{x}+50\frac{1}{x^2}+90\frac{1}{x^3}+20\frac{1}{x^4}$ 除以 $2x^3+5x+5+10\frac{1}{x}$.(他的答案是 $10x^3+x^2+4x+10+8\frac{1}{x^2}+2\frac{1}{x^3}$.)

12. 给出对结果

$$\sum_{i=1}^{n}i^3=\Big(\sum_{i=1}^{n}i\Big)^2$$

的完整的归纳证明,并与凯拉吉的证明相比较.

13. 运用伊本·海塞姆的步骤推导出整数五次幂和的公式:

$$1^5+2^5+\cdots+n^5=\frac{1}{6}n^6+\frac{1}{2}n^5+\frac{5}{12}n^4-\frac{1}{12}n^2.$$

14. 利用四次幂和与平方和的公式证明

$$\sum_{i=1}^{n-1}(n^4-2n^2i^2+i^4)=\frac{8}{15}(n-1)n^4+\frac{1}{30}n^4-\frac{1}{30}n$$
$$=\frac{8}{15}n\cdot n^4-\frac{1}{2}n^4-\frac{1}{30}n.$$

15. 证明可以通过双曲线 $y^2-x^2+\frac{d}{c}x=0$ 与抛物线 $x^2=\sqrt{c}y$ 的交来解方程 $x^3+d=cx$.画出这两条圆锥曲线.求出那些 c 和 d 值的集合,使这两条圆锥曲线分别为不相交,相交一次和相交两次.

16. 证明可以通过双曲线 $xy=d$ 和抛物线 $y^2+dx-db=0$ 的交来解方程 $x^3+d=bx^2$.假设 $\sqrt[3]{d}<b$,定出使这两圆锥线有零个,一个或两个交点的 b 和 d 所满足的条件.比较你的答案与萨拉夫·丁·图西对同一问题的分析.

17. 证明 $x^3+cx=bx^2+d$ 是奥玛·海亚姆的三次方程中惟一一个具有三个正解的方程.在什么条件下这三个解确实存在?

18. 用微积分证明 $x_0=2b/3$ 使函数 $x^2(b-x)$ 取极大值.然后用微积分分析 $y=x^3-bx^2+d$ 的图象,并确认萨拉夫·丁关于 $x^3+d=bx^2$ 的正解个数的结论.

19. 像萨拉夫·丁·图西所做那样,证明如果 x_2 是三次方程 $x^3+d=bx^2$ 的较大的一个正根,又如果 Y 是方程 $x^2+(b-x_2)x=x_2(b-x_2)$ 的正解,则 $x_1=Y+b-x_2$ 是原来的三次方程的较小的正根.

20. 分析 $x^3+d=cx$ 有正解的可能性,方法是先证明函数 $x(c-x^2)$ 的极大值发生在 $x_0=\sqrt{\frac{c}{3}}$ 处.利用微积分考虑 $y=x^3-cx+d$ 的图形并确定它具有零个,一个或两个正解时系数应满足的条件.

组合与数论方面的问题

21. 运用伊本·库拉定理证明 17 296 与 18 416 是亲和数.

22. 证明 1184 与 1210 是亲和数,但它们并非是塔比·伊本·库拉定理的结果.

23. 求一对与正文中不同的亲和数.(提示:试一试伊本·库拉定理的 $n=7$ 的情形.)

24. 用对 k 的归纳,给出伊本·班纳的结果 $C_k^n=\frac{n-(k-1)}{k}C_{k-1}^n$ 的一个现代证明.

几何方面的问题

25. 在假定伊本・海塞姆引理为真的条件下证明欧几里得公设5(平行公设),这个引理是:如果在一固定直线两端作两条夹角为直角的两条直线,那么从其中一条直线向另一条直线所引垂线等于这条固定的直线.
26. 证明下列的等式,这些等式是出现在对《原本》X 的评注者们著作中的典型材料:

(a) $\sqrt{\sqrt{8}\pm\sqrt{6}}=\sqrt[4]{4\frac{1}{2}}\pm\sqrt[4]{\frac{1}{2}}$.

(b) $\sqrt[4]{12}\pm\sqrt[4]{3}=\sqrt{\sqrt{27}\pm\sqrt{24}}=\sqrt[4]{51\pm\sqrt{2592}}$.

27. 阿布・萨赫勒・库西由他自己关于重心的工作以及他的先辈的工作知道,重心把某些平面和立体形状的轴分成比例如下:

三角形:$\frac{1}{3}$　　抛物线弓形:$\frac{2}{5}$　　四面体:$\frac{1}{4}$　　旋转抛物体:$\frac{2}{6}$　　半球:$\frac{3}{8}$.

在留意这些模型时,他猜想对半圆圈应是$\frac{3}{7}$.证明库西的前五个是正确的,但他对半圆圈的猜想暗含着 $\pi=3\frac{1}{9}$.(库西意识到这个值与阿基米德的上下界 $3\frac{10}{71}$ 与 $3\frac{1}{7}$ 相矛盾.但他的结论是,这是阿基米德著作的传播中的一个错误.)

三角学方面的问题

28. 导出凯比西的确定不可及对象的高的三角公式.
29. 用比鲁尼的步骤决定罗马的"奇伯拉"(纬度 41°53′N,经度 12°30′E).
30. 用比鲁尼的算法写出一个计算机程序来计算任何一个位置的"奇伯拉",其中这位置的经度和纬度已给定.用这个程序算出你所在地的"奇伯拉".
31. 比鲁尼设计了一个确定地球半径 r 的方法:从已知高度为 h 的山顶上观看地平线.这就是说,比鲁尼假设一个人可以测量出他的视平线与地平线的夹角 α(图 7.29).证明 r 由下面公式决定:

$$r=\frac{h\cos\alpha}{1-\cos\alpha}.$$

比鲁尼在一个特定情形下进行了测量,定出 $\alpha=0°34'$,这是从高度为 652;3,18 肘尺的山顶上量得的.并以肘尺为单位计算出地球的半径.假定一肘尺等于 18 英寸,把你的结果转换成英里,再与现代的数值相比较一下.对比鲁尼的做法的效力作出评论.

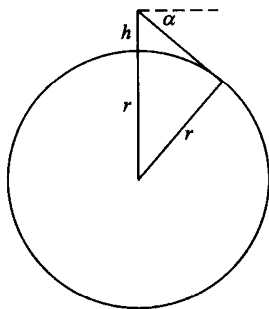


图 7.29　比鲁尼计算地球半径的方法.

32. 纳西尔・丁・图西在所有三条边都知道的情形下证明了一个解球面三角形的方法.设 ABC 为给出了弧 AB, AC 和 BC 的三角形(图 7.30).延长 AB 和 AC 分别至四分之一大圆 AD 和 AE.然后,通过 DE 作大圆并延长至 F.因为在 D 和 E 的角为直角,则四量规则表明 $\sin CF:\sin BF=\sin CE:\sin BD$.因为已知 $CE=90°-CA$ 和 $BD=90°-BA$,故而知道了 CF 的正弦与 BF 的正弦的比.同时可知 $BF-CF=BC$.证明在这些条件下,可以求出 BF 和 CF 弧.然后证明由球面直角三角形的结果可以求出 DF 和 EF 弧,因而可以求出角 $A(=DE$ 弧).剩下的那些角则可由正弦法则求出.应用此法解各边为 60°, 75° 和 31° 的球面三角形.
33. 按照正文中所表述的方法计算 $y=60\sin 1°$ 的近似值到60进制的头四位数.你的计算应该表明为什么迭代法有效.

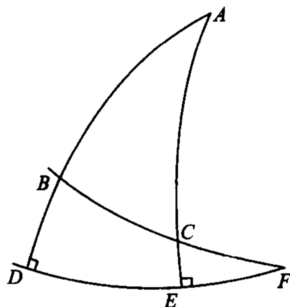


图 7.30　解出所有边已知的球面三角形.

讨论题

34. 为什么从十进位值制系统引进到它成为伊斯兰世界广泛使用的数字系统竟花了许多个世纪?
35. 设计一堂以花拉子米风格的几何推理教授二次方程求解公式的课.
36. 比较和对照花拉子米和塔比·伊本·库拉的二次方程求解公式的几何证明.哪一种方法更易于解释.
37. 设计一堂微积分方面的课,使用萨拉夫·丁·图西的方法以证明各种三次方程解的个数.
38. 设计一堂按照班纳的工作推导出 C_k^n 的乘法公式的课.
39. 设计一堂三角学的课,内容是应用解球面三角形的规则于各种令人感兴趣的问题.
40. 应该在今日的中学中教球面三角吗?给出赞成和反对的理由,但首先请确定一下,在哪些国家还在教此课程.
41. 在已知伊本·海塞姆的确定旋转抛物体体积的"积分"法和他确定整数的 k 次幂和的一般规则之后,为什么伊斯兰数学家没有发现对任意正整数 n 曲线 $y = x^n$ 下方的面积为 $\frac{x^n}{n+1}$?使伊斯兰数学家发现微积分则需要在伊斯兰文明中发生什么情况?

文献和注解

关于伊斯兰数学的一般著作中最好的是 Adolf P. Youschkevitch 著的 *Les Mathématiques Arabes*(VIIIe – XVe *siècles*)(Paris; J. Vrin. 1976).这是一本更广泛的关于中世纪数学的著作(俄文)中的一部分,它被 M. Cazenave 和 K. Jaouiche 翻译成了法文.另一本处理伊斯兰数学的各种重要思想的优秀英文书是 J. Lennart Berggren 的 *Episodes in the Mathematics of Medieval Islam*(New York: Springer-Verlag, 1986).这本著作虽然其本身而言并不是一本伊斯兰数学史,但却处理了伊斯兰数学家曾考虑过的某些重要的数学思想,而且是以大学数学学生能接受的水平撰写的. Roshdi Rashed 的一本论文集最近已编辑出版并被翻译为英文,书名为 *The Development of Arabic Mathematics: Between Arithmetic and Algebra*(Dordrecht: Kluwer, 1994).这个集子呈现了 Rashed 的许多最新的研究并提出了对伊斯兰数学的算术与代数方面的新观点.关于伊斯兰科学的一般综述(其中也包括了数学方面的材料)是由 Edward S. Kennedy 撰写的 The Arabic Heritage in the Exact Sciences,刊登在 *Al-Abhath* 23(1970),327 – 344 上.这篇文章及其他一些关于伊斯兰科学方面的有用的文章都可在 D. A. King 和 M. H. Kennedy 编辑的书 *Studies in the Islamic Exact Sciences*(Beirut: American University of Beirut Press, 1983) 中找到.

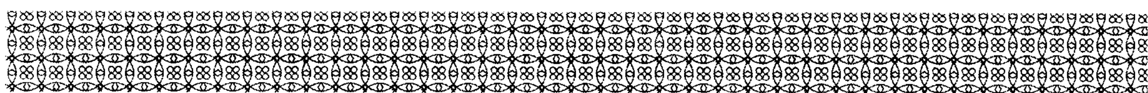
1. Jens Høyrup, "The Formation of 'Islamic Mathematics': Sources and Conditions," *Science in Context* 1(1987), 281 – 329, p. 306 – 307.
2. 同上.这篇文章包含了这些思想的详细叙述.作者认为:比较于希腊的情形,存在着一个伊斯兰"奇迹",它涉及到数学理论与实践的整合,而这对现代科学的创造是至关重要的.
3. A. S. Saidan, *The Arithmetic of Al-Uqlīdisī*(Boston: Reidel, 1978), p. 35.这本著作中提供了乌格里狄西著作的完整译文并配有评论.
4. 同上, p. 110.
5. 塞毛艾勒被引用于 Roshdi Rashed, *Arabic Mathematics*, p. 116.
6. 同上, p. 123.
7. Frederic Rosen, *The Algebra of Muhammed ben Musa*(London: Oriental Translation Fund, 1831), p. 3.这个从花拉子米的阿拉伯文《代数》的译文已长期脱印,很难找到.更近的从 12 世纪拉丁译本翻成的英文本是 Louis Karpinski, *Robert of Chester's Latin Translation of the Algebra of al-Khowarizmi*(Ann Arbor: University of Michigan Press, 1930).
8. 同上, p. 5.
9. 同上, p. 8.
10. Karpinski, *Robert of Chester's*, p. 69.

11. Rosen, *Algebra of*, pp.33 – 34.
12. 同上, p.12.
13. A. Sayili, *Logical Necessities in Mixed Equations by ' Abd al-Hamīd ibn Turk and the Algebra of his Time* (Ankara: Türk Tarih Kurumu Basimevi, 1962), p.166.
14. Rosen, *Algebra of*, p.22.
15. 同上, p.43.
16. 同上, p.54.
17. Martin Levey, *The Algebra of Abū Kāmil, Kitāb fi' lmuqābala, in a Commentary by Mordecai Finzi* (Madison: University of Wisconsin Press, 1966), p.32. 这是阿布·卡米勒著作的希伯莱文译本的英文编译, 并给出了他的代数与以前希腊和伊斯兰著作间关系的详细讨论.
18. 同上, p.144.
19. 同上, p.156.
20. Franz Woepcke, *Extrait du Fakhrī, traité d'algèbre par Aboù Bekr Mohammed ben Alhaçan al-Karkhī* (Paris: L'imprimerie Impériale, 1853), p.45. (这篇译文1982年由 Georg Olms Verlag, Hildesheim 重印) 请注意, 在这里凯拉吉的英文名字 al-Karajī 被写成 al-Karkhī. 不知道哪一个译名是正确的, 因为现存的阿拉伯文文献两者都有.
21. Adel Anbouba, *Al Samaw'al*, in the *Dictionary of Scientific Biography* (New York: Scribners, 1970—1980), vol. XII. 91 – 95; p.92.
22. Berggren, *Mathematics of Medieval*, p.114.
23. 关于伊本·海塞姆推理的更多细节参见 Roshdi Rashed, "Ibn al-Haytham et la mesure du paraboloide," *Journal for the History of Arabic Science* 5(1981), 262 – 291.
24. Rashed, *Arabic Mathematics*, p.65.
25. 同上. 也可见 J. Lennart Berggren, "Proof, Pedagogy, and the Practice of Mathematics in Medieval Islam," *Interchange* 21(1990), 36 – 48 and M. Yadegari, "The Binomial Theorem: A Widespread Concept in Medieval Islamic Mathematics," *Historia Mathematica* 7(1980), 401 – 406. Berggren不相信塞毛艾勒曾叙述过最一般的二项式定理, 因为完全没有他这样做的道理, 因此不可能说他已证明了它.
26. Daoud S. Kasir, *The Algebra of Omar Khayyam* (New York: Columbia Teachers College, 1931), p.49. 这本奥玛·海亚姆代数著作的英译本还包括了他对此学科贡献的详细讨论. 更多的关于奥马尔·海亚姆的讨论参见 D. J. Struik, "Omar Khayyam, Mathematician," *Mathematics Teacher* 51(1958), 280 – 285 and B. Lumpkin, "A Mathematics Club Project from Omar Khayyam," *Mathematics Teacher* 71(1978), 740 – 744.
27. 同上, p.44.
28. 关于萨拉夫·丁·图西更多的细节见 J. Lennart Berggren, "Innovation and Tradition in Sharaf al Dīn al-Tūsī's *al Mu' ādalāt*," *Journal of the American Oriental Society* 110(1990), 304 – 309, Jan P. Hogendijk, "Sharafal-Dīn al-Tūsī on the Number of Positive Roots of Cubic Equations," *Historia Mathematica* 16(1989). 69 – 85, and Roshdi Rashed *Arabic Mathematics*, Chapter 3. 萨拉夫·丁·图西的教科书被 Roshdi Rashed 翻译成法文, *Sharaf al-Dīn al-Tūsī, oeuvres mathématiques: Algèbre et géométrie au XIIe siècle* (Paris: Société d'Édition Les Belles Lettres, 1986). Rashed 不仅编译了这部阿拉伯文著作, 而且还给出了对其中所涉及的数学的广泛评述, 并特别联系到各种现代观念.
29. Ahmed Djebbar, *L'Analyse Combinatoire au Maghreb: L'Exemple d'Ibn Mun'im (XIIe-XIIIe) s.* (Orsay: Université de Paris Sud, Publications Mathématiques D'Orsay, 1985), pp.51 – 52. 这本小册子极详细地讨论了伊本·穆恩依姆的工作.
30. 同上, pp.55 – 56.
31. Rashed, *Arabic Mathematics*, p.300.
32. Rosen, *Algebra of*, pp.71 – 72. 花拉子米的《几何》的另一个英文译本是 Solomon Gandz, "The Mishnat Ha Middot and the Geometry of Muhammed ibn Musa al-Khowarizmi," reprinted in Solomon Gandz, *Studies in Hebrew Astronomy and Mathematics* (New York: Ktav, 1970). Gandz 强烈主张花拉子米对希伯莱几何的依赖性.
33. 同上, p.76.

34. 关于伊斯兰世界中的非欧几何的更多信息见 B. A. Rosenfeld, *A History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*, translated by Abe Shenitzer(New York: Springer-Verlag, 1988) 的第 2 章和 Jeremy Gray, *Ideas of Space: Euclidean, Non-Euclidean, and Relativistic*, second edition(Oxford: Clarendon Press, 1989) 的第 3 章. 也参见 D. E. Smith, "Euclid, Omar Khayyam, and Saccheri", *Scripta Mathematica* 3(1935), 5 – 10.
35. Ibn al-Baghdādī 被引用于 Galina Matvievskaya. "The Theory of Quadratic Irrationals in Medieval Oriental Mathematics," in D. A. King and G. Saliba, eds., *From Deferent to Equant: A Volume of Studies in the History of Science in the Ancient and Medieval Near East in Honor of E. S. Kennedy*(New York: New York Academy of Sciences, 1987)253 – 277, p. 267.
36. E. S. Kennedy, *The Exhaustive Treatise on Shadows by Abū al-Rayhān al-Bīrūnī*(Aleppo: University of Aleppo, 1976). p. 64. 这本书给出了比鲁尼著作的译文和对它的详细评论. 也可参见 E. S. Kennedy, "An Overview of the History of Trigonometry," in *Historical Topics for the Mathematics Classroom*(Reston, Va.: National Council of Teachers of Mathematics, 1989), 333 – 359.
37. 同上., p. 89.
38. 同上., p. 90.
39. 同上., p. 255.
40. 更详细的讨论见 Berggren, *Episodes in the Mathematics of Medieval Islam*, Section 6.8.

中世纪伊斯兰数学概览

780—850	花拉子米(Muhammad al-Khwārizmī)	算术、代数、实用几何
9 世纪	吐克('Abd al-Hamīd ibn Turk)	二次方程
830—890	库拉(Thābit ibn Qurra)	代数、亲和数
850—930	阿布·卡米勒(Abu Kāmil ibn Aslam)	使用无理数的代数
10 世纪中期	库西(Abu Ṣahl al-Kūhī)	重心
10 世纪中期	乌格里狄西(Abu l'Hasan al-Uqlīdisī)	最早的阿拉伯算术
940—997	阿布·瓦法(Muhammad Abū l-Wafā)	球面三角学的定理
11 世纪初	凯拉吉(Abu Bakr al-Karajī)	代数、归纳法最早的使用
11 世纪初	巴格达第(Ibn al-Baghdādī)	无理数
965—1040	海塞姆(Abu 'Ali ibn al-Haytham)	整幂的和、抛物体
973—1055	比鲁尼(Muhammad al-Bīrūnī)	三角学和应用
1048—1131	奥玛·海亚姆('Umar al-Khayyāmī)	三次方程、平行公设
1125—1174	塞毛艾勒(Ibn Yahyā al-Samaw'al)	十进制、多项式、二项式定理
12 世纪后期	图西(Sharaf al-Dīn al-Tūsī)	三次方程的分析
13 世纪初期	穆恩依姆(Ahmad ibn Mun'im)	组合与置换
1201—1274	纳西尔·丁(Nasīr al-Dīn al-Tūsī)	平行公设、三角学教科书
13 世纪后期	法雷西(Kamāl al-Din al-Fārisī)	亲和数和组合
1256—1321	班纳(Ahmad ibn al-Bannā)	组合结果的证明
15 世纪初	卡西(Ghiyāth al-Dīn al-Kāshi)	以十进制计算



第8章　中世纪的欧洲数学

> 要想学会正确测量和划分表面的方法，就必须透彻地理解讲授测量……所依赖的几何和算术中的一般定理. 如果一个人完全掌握了这些思想，他……就永远不会背离真理.
>
> ——引自《实用几何》(Liber Embadorum)，该书系[西班牙]萨瓦苏达(Savasorda)1116年希伯来文原著的拉丁文译本，由[意大利]蒂沃利的普拉托(Plato of Tivoli)翻译.[1]

神圣罗马皇帝腓特烈二世(Frederick II,1194—1250)爱好数学，正如我们从斐波那契(Fibonacci,Leonardo)1225年的一段记载中所知:"在被多米尼克老爷带回到比萨拜见陛下您，最辉煌的君主之后，我曾遇见巴勒莫的约翰老爷；他对我提出一个他所想到的问题，涉及算术与几何……当我最近得知尊贵的陛下您屈尊阅读了我所写的关于数的书[《算盘书》(the Liber abbaci)]，并喜欢了解一些关于几何与数字的技巧时，我想起您的科学随从在您的宫廷向我所提出的问题. 我便承担起这个论题，开始撰写本书，我欲称其为《平方数书》(The Book of Squares). 若其中含有不正确或不必要的东西，恳请您宽恕，因为能记住一切并从不犯错的是神，而非人类；没有人能避免错误，也没有人能处处周到."[2]

公元476年，西罗马帝国在各种蛮族的猛攻下灭亡. 在旧帝国的不同部分不久便形成一些封建社会，这意味着欧洲各民族国家的长期发展过程的肇始. 然而，在随后五个世纪中，欧洲文化的总体水平很低. 农奴耕作田地，贵族几乎都不会阅读或写作，更别提理解数学了. 确实，没有多少实际需要这个课题，因为封建集团相对是自给自足的，贸易几乎不存在，特别是穆斯林征服地中海沿岸之后.

尽管缺乏数学方面的努力，早期中世纪还是从古代继承了那种见解，认为应该要求受教育者学习四艺——算术，几何，音乐，及天文. 因此，圣奥古斯丁(St. Augustine,354—430)在他的《天堂》(City of God)中写道:"我们不应该鄙视数的科学，在《圣经》的许多节次中均可发现，它对那位

细致的解释者具有显著的帮助.如果没有道理,也不会用它来赞美上帝:'您用数字、大小及重量安排了世间万物.'"[3] 然而,那时用以研究"数的科学"的仅有课本是拉丁文版的尼科马科斯的《算术》,欧几里得的《原本》部分章节的编译本,以及音乐和天文的入门介绍,由罗马学者博伊西斯(Boethius,480—524)及塞维利亚主教伊西多雷(Isidore,560—636)编译和撰写.因而,数学四艺的框架还是有的,但仅是梗概而已.

当时仅有的学校实质上均与修道院有关,其中多数是来自爱尔兰的僧侣所建.爱尔兰是不属于神圣罗马帝国本土部分而又采用基督教的最早国家.当欧洲大陆的大部分地区处于动乱之时,爱尔兰修道院的僧侣复制了希腊和拉丁文的原稿,从而保存了许多古代学问.欧洲各地的学生汇聚那里去学习.然后,在6—8世纪间,教会派遣来自爱尔兰的耶稣会士到欧洲大陆建立学问中心,由此,新智慧的发展在几个世纪之后便出现了.

甚至在中世纪的最初阶段,有一个重大的数学问题要被考虑:确定历法.特别是教会要争辩,到底是应该用罗马的日历还是用犹太人的阴历来确定复活节.这两种方法体系只有通过数学知识才能一致.因此,查理大帝(Charlemagne)甚至在公元800年被加冕为神圣罗马帝国皇帝之前就正式建议,复活节计算所必需的数学应是教会学校课程的一部分.

查理大帝聘请英国约克的阿尔昆(Alcuin of York,735—804)做他的教育顾问,帮他建立更多的学校.阿尔昆曾与一名爱尔兰教师一起进行研究并在查理大帝的宫廷中有几名爱尔兰牧师相协助,在需要书籍时他通常便派人去英格兰和爱尔兰.我们没有多少关于他的数学知识的直接材料,唯有当时的一本关于53道算术题的问题集通常被认为是他的作品,书名为《敏锐青年之命题》(Propositiones ad acuendos juvenes).该习题集的题目常常需要一些机敏才能解决,但并不依赖任何数学理论或程序法则.

公元10世纪,数学兴趣的复活始于奥里亚克的热尔贝(Gerbert d'Aurillac,945—1003)的工作,他在999年成为罗马教皇,改称西尔维斯特二世(Pope Sylvester II)(图8.1).在青年时代,热尔贝曾在西班牙学习,在那里可能学过一些穆斯林的数学.后来,在神圣罗马皇帝奥托二世(Otto II)的庇护下,热尔贝重组了兰斯(Reims)地方大教堂的学校,并重新引入数学学习.除了讲授初等算术和几何,热尔贝曾涉及过罗马测量员的测量法则和天文基础.他还教过算盘的使用,算盘被分成不同的档表示十的不同次幂,在每一档放置一个筹码并用阿拉伯数码1,2,3,…,9之一注明,并用空档表示零.热尔贝的工作标志着印度-阿拉伯数字在基督教西方的最早出现,但其中没有包含零且缺乏恰当的利用这些筹码进行计算的法则,这表明热尔贝并未理解印度-阿拉伯数码系统的全部含义.



图8.1　奥里亚克的热尔贝,即罗马教皇西尔维斯特二世.

虽然千年之交时提供给欧洲人的原始著作很有限,学者们还是知道希腊数学家已经取得许多成就,不过那时他们实际上难以看到.这种遗产,以及伊斯兰教世界所发展的那部分数学,通过译作才被带入西欧.欧洲学者发现较大范围的希腊科学著作(主要在阿拉伯语言译本中)始于12世纪,并开始了将其译成拉丁文的进程.大部分工作在西班牙的托莱多(Toledo)完成,那些著作是耶稣会士从以前的穆斯林统治者那里取回的.其中可发现大量伊斯兰科学手稿,并了解到许多跨越两种文化的人士.特别地,有一个繁荣的犹太人社会,其中许多成员精通阿拉伯语言.翻译通常采取两个阶段,先由西班牙的犹太人从阿拉伯语译成西班牙语,再由基督教学者从西班牙语译成拉丁语.下面的主要数学著作译本一览表及翻译日期还是比较详尽的(补遗8.1).

补遗 8.1	翻译家和他们的译著
	威尼斯的詹姆斯(约1128—1136)
	亚里士多德的《正位篇》,《前分析篇》,《后分析篇》.
	巴思的阿德拉德(约1116—1142)
	花拉子米的天文表
	欧几里得《原本》
	花拉子米的算术著作(Liber ysagogarum Alchorismi)
	塞维利亚的约翰和多明哥·冈迪萨尔沃(约1135—1153)
	花拉子米算术著作编译(Liber alghoarismi de practica arismetrice)
	蒂沃利的普拉托(约1134—1145)
	奥多修斯(约100 B.C.)的《球面学》(Spherica)
	阿尔·巴塔尼的《星的运动》(De motu stellarum),其中包含重要的三角学材料
	阿基米德的《圆的度量》
	萨瓦苏达的《实用几何学》
	切斯特的罗伯特(活跃于1141—1150)
	花拉子米的《代数学》
	修订花拉子米的天文表并用于伦敦子午线确定
	克雷莫纳的杰拉德(活跃于1150—1185)
	亚里士多德的《后分析篇》
	奥托利可斯的球面学著作(De sphaera mota)
	欧几里得的《原本》
	欧几里得的《数据》
	阿基米德的《圆的度量》
	西奥多修斯的《球面学》
	托勒密的《大成》
	梅内劳斯的《球面》
	花拉子米的《代数学》
	贾比尔的《天文基础》(Elementa astronomica)
	莫贝克的威廉(Wilhelm of Moerbeke,活跃于1260—1280)
	阿基米德的《论螺线》
	阿基米德的《论平面圆形的平衡》
	阿基米德的《抛物线求积》
	阿基米德的《圆的度量》
	阿基米德的《论球和圆柱》
	阿基米德的《论劈椎曲面和旋转椭球》
	阿基米德的《论浮体》
	注:本表所含原著,其译本的译者均可确认.另有一些原著,其译本出现于12和13世纪,但译者不详,如阿波罗尼奥斯的《圆锥曲线论》部分章节和艾布卡米勒的《代数》.

在早期的翻译成员中有[西班牙]塞维利亚的约翰(John)和多明哥·冈迪萨尔沃(Domingo Gundisalvo),他们在12世纪上半叶是非常活跃的.约翰出身于犹太人,但皈依了基督教,他的原名可能是所罗门·本·大卫(Solomon ben David),而冈迪萨尔沃是一位哲学家和基督教神学家.他们

最重要的数学译作是根据花拉子米的算术成就写成的著作《花拉子米实用算术书》.他们也翻译了大量天文著作,包括对托勒密著作的注解,及许多医学和哲学著作.

一位与塞维利亚的约翰同时代的人是巴思的阿德拉德(Adelard,1075—1164),他生于巴思,青年时代曾花费大量时间游历法国,南意大利,西西里,和近东.特别在后两个地区可接触到大量阿拉伯语的专著.阿德拉德是最早将阿拉伯语的欧几里得《原本》翻译过来的人.他还在1126年翻译过花拉子米的天文表.该译本含有以拉丁文出现的最早的正弦表和正切表,而后者是一位11世纪的编者添加到花拉子米著作中的.另一位英国人是切斯特的罗伯特(Robert),曾在西班牙住过几年,在1145年翻译了花拉子米的《代数》,因而在欧洲传播了解二次方程的代数算法.在同一年,蒂沃利的普拉托从希伯来文翻译了西班牙的犹太学者萨瓦苏达的《实用几何》,该书也包含了伊斯兰解二次方程的法则.

最多产的翻译家是克雷莫纳的杰拉德(Gerard,1114—1187),他是意大利人,长期工作在西班牙的托莱多,被认为完成了80多部著作的翻译.无疑,并非所有著作都是他一人翻译.大多数是在他的指导下由别人来做,但历史已遗忘了绝大多数他的助手的姓名.杰拉德的译著包括从塔比·伊本·库拉(Thābit ibn Qurra)的阿拉伯文译著翻译的欧几里得《原本》的新译本,和从阿拉伯文翻译的托勒密的《大成》(Almagest)的最早译本(1175年).

到12世纪末,欧洲能阅读拉丁文的学者均可涉猎大多数希腊数学的主要著作.在随后几个世纪中,这些著作被吸收,欧洲人自己开始创造新数学.然而,应该注意一些西班牙的犹太学者阅读阿拉伯文原著要早些,那时已经用希伯来文独立地撰写著作了.事实上,在12世纪期间,欧洲和地中海流域的三种文明(犹太人的,基督教的和伊斯兰教的)之间的文化交流是非常密切的.但伊斯兰前几个世纪的至尊日益衰落,其它两种文明正在积蓄力量.到13世纪末,西方基督教世界的才华则显露出来,而对犹太人环境的各种限制减少了犹太人的贡献.

本章将讨论12世纪到14世纪之间犹太人及基督教的贡献.首先将考虑几何学和三角学,其次是组合学的发展,然后是伊斯兰代数传入欧洲而产生的代数学,最后是运动的数学,其来源是中世纪大学对亚里士多德著作的研究.

8.1 几何学和三角学

欧几里得《原本》在12世纪早期被译成拉丁文.当然之前西班牙已有阿拉伯文的版本.因此,当巴塞罗那的萨瓦苏达(卒于1136年)在1116年撰写他的希伯来文《实用几何》(Hibbur ha-Meshihah ve-ha-Tishboret)、以帮助法国和西班牙的犹太人测量他们的田地时,在他的著作的开始,便有一个关于欧几里得的一些重要定义、公理和定理的摘要.对萨瓦苏达的生平所知不多,他本名阿勃拉汉·巴·希亚(Abraham bar Hiyya),萨瓦苏达(Savasorda)是他的拉丁文称号,是阿拉伯词语"卫队长"的讹译,由此来看,很可能他曾有过一个宫廷职位,也许是基督教君王的数学和天文顾问.

8.1.1 萨瓦苏达的《实用几何》

像随后几个世纪的绝大多数几何学家一样,萨瓦苏达感兴趣的是几何方法对测量的实际应用,而非欧氏几何《原本》的理论方面.不过,他采用了从希腊人那里吸收来的伊斯兰的证明传统,并就解代数问题的方法给出几何验证,作为书中几何讨论的一部分.特别地,他将《原本》卷II中关于几何代数的主要结论包含在他的著作中,并用他们来证实解二次方程的方法.萨瓦苏达的著作在欧洲

最早给出了解这种方程的伊斯兰算法.

例如,萨瓦苏达提出这样的问题:“如果从一正方形的面积中减去四边之和剩余 21,那么,正方形的面积和每一等边的长各是多少?”[4] 我们可将萨瓦苏达的问题改写成二次方程 $x^2 - 4x = 21$. 他是用熟知的方式解方程的:4 的一半得 2,将这个结果平方得 4,将这个平方加到 21 得 25,求平方根得 5,然后将其加到 4 的一半便得到边长的答案 7,面积的答案 49. 在问题的陈述中用面积减长度(四边之和),这并非几何语言. 但在几何验证中,他把问题重述为,从未知边是 x 的原正方形中减去边是 4 和 x 的长方形,剩余的长方形面积为 21. 然后,他二等分长为 4 的边,并用《原本》卷 Ⅱ 命题 6 验证这个代数步骤. 这明显地说明萨瓦苏达所学代数并非源于花拉子米,而是来自艾布 · 卡米勒(Abū Kāmil),艾布 · 卡米勒曾使用过欧几里得验证法. 萨瓦苏达对其它两种伊斯兰的混合二次方程也提出过类似的解法和欧几里得证明,例如,$x^2 + 4x = 77$ 和 $4x - x^2 = 3$. 在后一情况中,他给出了两个正数解. 萨瓦苏达还解了如下的二次方程组 $x^2 + y^2 = 100, x - y = 2$ 和 $xy = 48, x + y = 14$.

然而,萨瓦苏达的原创性工作是在他的测圆部分. 他从给出求圆周和圆面积的标准法则开始,首先用 $3\frac{1}{7}$ 代表 π,然后指出如果想得到更精确的值,比如在作天文计算时,就应该使用 $3\frac{8\ (1/2)}{60}$ $(= 3\frac{17}{120})$. 奇怪的是,在希伯来文版本中有一处是用无穷小方法验证面积公式 $A = \frac{C}{2}\frac{d}{2}$(见图 1.12),而在拉丁文版本中却没有这样的验证. 要测量圆的弓形面积,萨瓦苏达解释道,就应该用半径乘以弧长的一半求出相应的扇形面积(图 8.2). 然后,减去由该弧线所对应的弦及其端点处的两个半径所构成的三角形的面积. 但是,假定知道弦长,如何求其弧长呢?可以使用弧弦表. 这是欧洲第一个可以称之为三角表的东西(图 8.3). 它不同于花拉子米的正弦表,花拉子米的正弦表的拉丁译本出现在著作之后,而且是用度数测量弧长的,圆的半径是 60,而萨瓦苏达的弧弦表是用更便于测定的弦长求弧长的. 即,他分半径为 14 段(parts),因而周长之半是整数 44,然后,对应于从 1 到 28 的每一个整数值的弦长,以段,分,秒为单位给出其弧长. 给定弦长 S 和从弦的中心到圆周的距离 h 后,为了确定该弓形的弧长,萨瓦苏达先用下面的公式确定圆的直径 d(图 8.4):

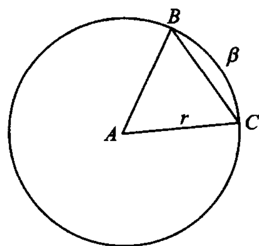


图 8.2　弓形 $B\beta C$ 的面积 = 扇形 $AB\beta C$ 的面积 − 三角形 ABC 的面积;扇形面积 $= r \cdot \frac{\beta}{2}$.

$$d = \frac{S^2}{4h} + h.$$

然后,他用 $28/d$ 乘以弦长 S(转换到直径是 28 的圆),查阅他的表确定相应的弧 α,再用 $d/28$ 乘以 α 求出实际弧长. 萨瓦苏达的上述步骤提供了三角表被应用于测量地上的而不是天上的量的最早实例之一. 但要注意,在这一情形三角学是被用来测定圆的切片[弓形]而非三角形的.

8.1.2　实用几何

萨瓦苏达的希伯来文原著是中世纪欧洲的许多早期实用几何著作之一. 另一本拉丁文著作,出现在 1120 年代,也许是圣维克多的于格(Hugh,1096—1141)撰写的,他是一位神学家和巴黎圣维克多修道院的主持. 该书对象是测量员,要比萨瓦苏达的著作简单得多. 显而易见,三角学知识还没有传到巴黎,书中也没有提到欧几里得. 然而,于格使用了星盘,这种观测仪由伊斯兰天文学家从早期的希腊模型发展而来,并经西班牙传入西欧. 因而,于格的测量法就涉及到照准仪的使用,照准仪是

Partes Cordatum	Arcus		
	Partes	Min.	Sec.
1	1	0	2
2	2	0	8
3	3	0	26
4	4	0	55
5	5	1	44
6	6	2	54
7	7	4	42
8	8	7	11
9	9	9	56
10	10	13	42
11	11	18	54
12	12	24	38
13	13	31	9
14	14	40	0
15	15	50	10
16	17	2	16
17	18	16	36
18	19	33	27
19	20	53	26
20	22	17	10
21	23	45	6
22	25	19	24
23	27	0	0
24	28	49	56
25	31	26	37
26	33	20	52
27	36	27	32
28	44	0	0

图 8.3 萨瓦苏达的弧弦表.

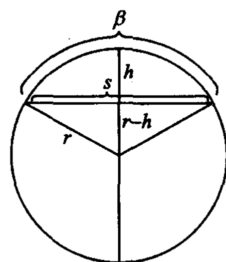


图 8.4 弧长 $\beta = \frac{d}{28}$ · 弦$(\frac{28}{d}s)$所对应的弧,其中 $d = 2r = \frac{s^2}{4h} + h$.

附属于星盘的测高装置,能测定观测物的高度与距离之比率(图 8.5).如果已知该比率 r,且观测物的距离也是知道的,那么高度 h 就可用 $h = rd$ 确定.像印度,中国,和伊斯兰世界的前辈那样,于格知道并非总能测出远方物体的距离.在这种情况下,就需要两次测量(图 8.6).在点 S_1,得出高度 h 与距离 d_1 之比 r_1,而在点 S_2,得出 h 与 d_2 之比 r_2.由此知 $d_2 = (r_1/r_2)d_1$.但因 $d_2 - d_1 = f$ 能被测出,故于格便算出 d_1 为

$$d_1 = \frac{f}{\frac{r_1}{r_2} - 1},$$

并用 $h = r_1 d_1$ 估算 h.[5]

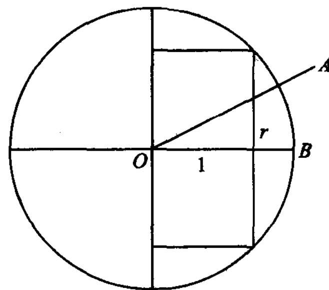


图 8.5 装有照准仪 OA 的星盘.将线 OB 定为地平线并沿 OA 瞄准观测物.则 r 为观测物的高度与距离之比率.

然而,到 12 世纪后期,三角学和欧几里得的知识已经传入巴黎,这可在一本无名的实用几何著作中得到例证,该书通常依据手稿的前三个词 Artis cuiuslibet consummatio 而被叫做《艺术至善》.这本原来用拉丁文写成,但在 13 世纪译成法文的著作开始有一段富有诗意的介绍:

艺术至善,总体来看,取决于两方面:理论与实践.仅知其一,难称完美.实际上当代的

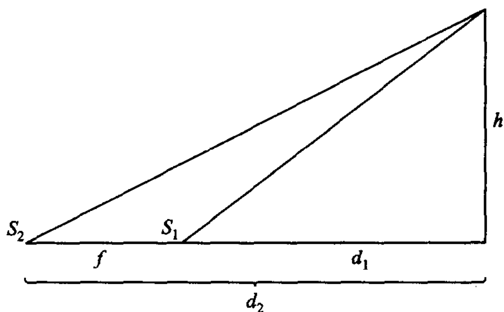


图 8.6 圣维克多的于格用两次观测法测量远方物体的高度.

拉丁语系人…… 由于忽视实践,在辛勤播种的地方却无所获,犹如未待其结果便采摘春天的花卉.一旦能用算术推知数字的特性,而非通过巧算认识它们的层出不穷,那将多么优美;一旦能用音乐推知音律,而非通过听觉辨识它们的和谐,那将更加动听悦耳;一旦能用几何证出边长和角度,而非直接测知它们的大小,那将多么神妙;一旦能用天文学推知星球的运动,而非仅看到日月食和天文奥秘,那将多么辉煌和卓越.因此,我们准备了这本实用几何学论著,以便将我们从大师的原著中所汲取的知识提供给有兴趣的读者.[6]

作者似乎想说明,要真正得到教育,就不应该仅研究四艺的理论方面,还须了解如何将这些课题用于现实世界.那么,《艺术至善》就是要展示四艺之一,即几何学的实际应用.

该书有四个部分:面积测量,高度测量,体积测量,分数计算.最后一部分为前面各部分提供所必要的计算知识.关于面积的第一部分,开始是一些求三角形,矩形,和平行四边形面积的基本过程,其中大多数可用欧几里得命题进行验证.随后是各种等边多边形的面积.但所列公式都不正确.并不是给出边长是 n 的五边形,六边形,七边形,…… 的面积公式,而是给出了求第 n 个五边形,六边形,七边形数的公式.例如,边长为 n 的五边形面积公式相当于

$$A = \frac{3n^2 - n}{2}.$$

作者可能深受形数材料的影响,而形数起源于尼科马科斯的著作.

关于高度的那部分则显示出作者的三角学知识.例如,通过一个长度为 12 的垂直标杆的影长,测量太阳高度的过程如下:"把影长自乘并加上 144,求其平方根.然后,用 60 乘以影长,得到的乘积再除以所求出的平方根.结果便是正弦值;求其弧度,并从 90 中减去该弧度;其差便是 …… 太阳的高度".[7] 亦即,如果影长为 s,那么高度 α 为

$$\alpha = 90 - \arcsin\left(\frac{60s}{\sqrt{s^2 + 144}}\right),$$

注意,正像过去大多数伊斯兰的三角学著作一样,正弦值的计算,使用的是半径为 60 的圆(图 8.7).类似地,作者也给出了从高度求影长的公式

$$s = \frac{12\sin(90 - \alpha)}{\sin \alpha}.$$

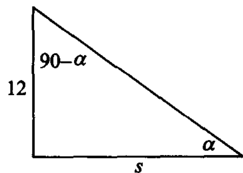


图 8.7 《艺术至善》中用影长求日高及其逆问题:
$\alpha = 90 - \arcsin\left(\frac{60s}{\sqrt{s^2 + 144}}\right)$;
$s = \frac{12\sin(90 - \alpha)}{\sin \alpha}$.

上述两个问题说明,作者知道正弦表的使用,但也许不知道余弦,正切,或余切表,而这些表都已经出现在当时的伊斯兰世界.当时只能接触到印度和伊斯兰人对希腊成果的一些早期改进.

对于观测,作者则回到古代的方法.如要测塔高,他不仅不用三角学方

法，而且回到了也许是最古老(和最简单)的方法："等到太阳是45度时……；那么，任何物体在平面上的影长就等于该物体."[8] 如果塔不可接近，作者则使用古代的两次观望法，即类似于中国和印度资料中的方法.从几乎所有的印度，伊斯兰，和中世纪欧洲的资料来看，甚至在已经掌握了三角学方法后，它们也仅仅是被应用于天文，而非日常生活.

这两本12世纪的拉丁文几何学给了我们关于那个时代北欧几何学知识的一个整体概念.希腊的几何学传统刚开始被重建，起源于古代并非全部都严格正确的实用几何学方法，被用来计算日常生活中的几何量.然而，在南欧，伊斯兰的影响更强大，欧几里得的证明传统更为明显，萨瓦苏达的著作就是例证.另一个例证是早期的意大利数学家之一，比萨的莱昂那多(约1170—1240).

8.1.3　比萨的莱昂那多的《实用几何学》

就与《艺术至善》或圣维克多的于格的著作的关系而言，莱昂那多的《实用几何学》(Practica geometriae，1220)更接近于萨瓦苏达的著作.事实上，某些段落似乎就直接取自于萨瓦苏达的《实用几何学》，然而，莱昂那多的著作相对来说更广泛详细.同早期著作一样，莱昂那多开始先列举了若干欧几里得的定义，公理，和定理，特别是包含了卷II的命题.因此在测量矩形一节中(其中含有解二次方程的标准方法)，他就能够引用欧几里得来验证他的算法.相比萨瓦苏达而言，他提供了更多的例子，包括平方项的系数大于1的方程.例如，为解方程 $3x^2+4x=279$，他先除以3化简为 $x^2+1\frac{1}{3}x=93$，再应用标准方法.另外，他的许多问题涉及矩形的对角线，因此，就论述了边长的平方和.

莱昂那多又同萨瓦苏达一样，写了关于圆的一节，其中引用了 π 的标准近似值22/7.但是，莱昂那多重新展示了如何利用阿基米德的方法计算这个值.他得出圆的外切正96边形的周长与圆的直径之比是 $1440:458\frac{1}{5}$，以及圆的内接正96边形的周长与圆的直径之比是 $1440:458\frac{4}{9}$.注意到 $458\frac{1}{3}$ 是 $458\frac{1}{5}$ 与 $458\frac{4}{9}$ 的近似平均值，他便断言圆周与直径之比近似于 $1440:458\frac{1}{3}=864:275$.因为 $864:274\frac{10}{11}=3\frac{1}{7}:1$，莱昂那多便重新导出了这个阿基米德的值.

莱昂那多还计算了圆的弓形和扇形的面积.为此，他需要一个弧弦表.奇怪的是，虽然他按照标准方式定义了弧的正弦，却没有给出正弦表，而是一个弦表，并且他重新介绍了托勒密从整弧的弦确定半弧的弦的方法步骤.他的弦表不同于托勒密的正弦表.事实上，因其基于半径是21的圆，很可能就是莱昂那多的原创.正像萨瓦苏达取值14一样，选择21是为了使圆的周长之半为整数，但不同于萨瓦苏达的弧弦表，这个表是一个直弦表(图8.8).对于每一整数弧长从1到66(还有从67到131)，该表以相同的度量单位(rods)给出了相应的弦长，但该度量单位的分数部分并非60进制，而是比萨的度量单位feet(6 feet = 1 rod)，unciae(18 unciae = 1 foot)，和points(20 points = 1 uncia).莱昂那多然后展示了在半径不是21的圆中，如何利用这个弦表通过弧长计算弦长.

像萨瓦苏达一样，莱昂那多仅用弦表计算圆的扇形和弓形面积.在同一章的后面，当他计算圆的内接正五边形的边长和对角线时，他没有使用对我们来说似乎很明显的查阅弦表的方法.他返回到欧几里得，并从卷III中引用了一些有关六边形，五边形和十边形边长的定理，以便能让自己实际完成这种计算.直到著作的最后，当他想要计算海拔高度时，他仍然没有使用三角学.他使用的是过去的相似三角形的方法，先用已知高度的标杆帮助瞄准未知物的顶端，然后再测量出沿地面的相关距离.

Arcus pertice	Arcus pertice	Corde pertice	Ar pedes	Cuu vncie	M puncta	Arcus pertice	Arcus pertice	Corde pertice	Ar pedes	CV vncie	VM puncta
1	131	0	5	17	17	34	98	30	2	6	17
2	130	1	5	17	13	35	97	31	0	8	5
3	129	2	5	17	4+	36	96	31	4·	8	7
4	128	3	5	17	2	37	95	32	2	5	15
5	127	4	4	12	10	38	94	33	0	1	9
6	126	5	5	16	7+	39	93	34	3	13	0
7	125	6	5	14	5	40	92	35	1	4	15
8	124	7	5	12	9	41	91	35	4	12	10
9	123	8	5	8	16	42	90	36	2	0	0
10	122	9	5	7	8	43	89	36	5	3	5
11	121	10	5	4	2	44	88	37	2	4	6
12	120	11	4	17	18	45	87	37	6	3	2
13	119	12	4	13	6	46	86	38	1	17	15
14	119	13	4	7	16	47	85	38	4	12	13
16	117	14	4	1	0	48	84	38	1	4	0
16	116	15	3	11	18	49	83	38	3	11	15
17	115	16	3	3	12	50	82	39	5	17	2
18	114	17	2	12	8	51	81	40	2	2	1
19	113	18	2	0	15	52	80	40	4	2	10
20	112	19	1	8	12	53	79	40	0	0	11
21	111	20	0	13	18	54	78	40	1	14	5
22	110	21	0	0	0	55	77	41	3	7	8
23	109	21	5	2	16	56	76	41	4	16	2
24	108	22	4	4	5	57	75	41	0	4	12
25	107	23	3	4	8	58	74	41	1	8	1
26	106	24	2	3	2	59	73	41	2	9	0
27	105	25	1	0	6	60	72	41	3	7	14
28	104	25	5	10	2	61	71	41	4	9	2
29	103	26	4	8	0	62	70	41	4	15	10
30	102	27	3	0	3	63	69	41	5	6	9
31	101	28	1	9	7	64	68	41	5	12	17
32	100	28	5	16	4	65	67	41	5	6	14
33	99	29	4	3	9	66	66	42	0	0	0

图 8.8 比萨的莱昂那多的弦表.

人物小传	比萨的莱昂那多(约 1170—1240)(Leonardo of Pisa)
	莱昂那多生于 1170 年左右,现在通常被称为斐波那契(意为波那契之子),此名源于 19 世纪时其著作的编订者邦孔帕尼(B · Boncompagni). 父亲是比萨商人,在北非口岸今日阿尔及利亚的贝贾亚有大量商务交易. 莱昂那多早年在那里师从穆斯林教师学习阿拉伯语和数学. 后来他也许为了父亲的商务而游遍地中海地区. 每到一地,他都寻访伊斯兰学者并吸取伊斯兰世界的数学知识. 大约在 1200 年返回比萨后,他花费了 25 年时间将其所学陆续整编成书. 其中保存至今的有《算盘书》(1202,1228),《实用几何》(1220),和《平方数书》(1225). 正如本章开头的故事所述,莱昂那多的重要地位已得到腓特烈二世宫廷的认可. 不仅如此,比萨市在 1240 年曾授予他一次年俸以感谢他的贡献.

8.1.4 三角学

三角学在中世纪时期并不用于测量日常三角形,这可以在 14 世纪的两本三角学著作中得到进一步证实,其中之一的作者是英国沃灵福德的理查德(Richard,1291—1336),另一本的作者是法国犹太人莱维本热尔松(Levi ben Gerson,1288—1344).

沃灵福德的理查德是一位修道士,作为圣奥尔本斯(St. Albans) 修道院院长度过了人生的最后九年. 关于三角学基础的著作《四分集》(Quadripartium) 撰于他还是牛津大学的学生时代,即 1320 年左右. 约十年后,理查德在另一部名为《论分割》(De Sectore) 的著作中,修订和压缩了这部著作. 像那时的多数三角学课本一样,这两部著作的目的是讲授解球面三角学问题所必要的方法,这些方法也是天文学的需要.《四分集》的主要来源似乎是托勒密的《大成》,除较早的弦表外,修改增添了印度的正弦表. 但到理查德订正这部著作时,他已通晓贾比(Jābir) 的三角学. 事实上,在球面三角学一节中,他在提出了托勒密关于梅内劳斯(Menelaus) 定理的描述之后,又叙述了贾比以梅内劳斯定理为基础的所有论述.

理查德对梅内劳斯定理的论述,不管是平面还是球面的情形,都非常详细. 因为这个定理涉及到梅内劳斯位形中各种边长的比率,理查德需要首先考虑比例定理的基础. 他对比例的研究密切相关于大学里几位同时代人的著作,这将在 8.4.1 节中讨论. 这里我们仅指出,在对梅内劳斯定理的论述中,理查德讨论了梅内劳斯位形的所有可能情形,每次都重新给出证明. 也许现代读者认为他的著作冗长,但他明显地感到这对那些缺乏数学经验的读者是必要的,而他就是为这些读者在撰著. 人们可以在此处,还可以在该书的开始章节关于平面三角学的基本结果中看到,当理查德详尽无遗地论述所有情形时,所体现的他对严格的欧几里得论证的赞成和信奉. 从历史上看,虽然数学知识在中世纪早期处于衰败不振,但数学证明的基本思想还是幸存了下来,一旦对更多数学的需求确定之后,它们便会重新活跃起来,这一点在理查德的著作中得以反映.

莱维本热尔松的三角学著作几乎与《四分集》同时出现. 它是一部天文学专著的一部分,而这部天文学专著又构成一部名为圣战(Sefer milhamot Adonai) 的主要哲学著作的一部分. 莱维的三角学主要以托勒密为基础,并且也与理查德一样,莱维通常使用正弦而非弦长. 莱维与托勒密以及理查德的主要区别是他给出了解平面三角形的详细过程. 他先给出解直角三角形的标准方法,然后进入到一般三角形. 在已知三边的情况下,莱维通过从一顶点向对边(或对边的延长线) 作垂线,然后应用《原本》卷 II 的命题 12 和 13 中的余弦定理的论述,来解三角形. 这一方法也适用于已知两边及其夹角的情况. 在已知两边及其中之一的对角的情况下,莱维使用(并证明) 了正弦定理. 然而,他

没有提及这种情况可能蕴涵两个不确定的三角形.当然,在任一具体问题中,只要规定两未知角之一是锐角或钝角,就能得到惟一的三角形解.最后,莱维指出,已知两角和一边的情况也能用正弦定理解出.

无疑,莱维的方法不是新创的.虽然他的过程有点不同于贾比,但在伊斯兰的其它著作中还是可以见到的.不过,莱维的简短论文提供了早期欧洲关于解平面三角形的基本方法的论述之一.但正像伊斯兰的著作和实用几何课本一样,莱维所提出的这些方法仅用于解天文三角形,从未用于解日常生活中的三角形.

8.2 组合学

我们已经讨论过印度和伊斯兰资料中组合学方面的兴趣.中世纪欧洲对这样的问题也有兴趣,特别是在犹太人社会中.犹太人关于这个论题的最早资料似乎是那本神秘的著作《天地万物》(Sefer Yetsirah),该书写于8世纪前的某个时期,也许最早是2世纪.其中,那位无名作者计算了22个希伯来文字母的各种排列方式.他对这种计算感兴趣是因为犹太神秘主义者相信上帝以这种方式命名事物而创造了宇宙万物:"上帝抽取和组合字母,斟酌和交换字母,且用字母产生了整个宇宙和预定要创造的一切 …… 两个字母组成两个单词,三个字母组成六个单词,四个组成二十四个单词,五个组成一百二十个单词,六个组成七百二十个单词,七个组成五千零四十个单词."[9] 明显地,作者知道 n 个字母的可能排列数是 $n!$.一位意大利的犹太教教士,多诺洛(Shabbetai Donnolo,913—970),在关于《天地万物》的评注中非常明确地导出了这个事实上的规律:

> 由两个字母[不重复地]组成单词,第一个字母可置换两次.对于三个字母的单词的每一个打头字母来说,其它的字母能被置换形成两个由两个字母组成的单词来对应三次中的每一个打头字母.三个字母的单词的所有排列对应于由四个字母组成的单词的每一个打头字母:三个字母的单词有六种方式,因此对于四个字母的单词的每一打头字母有六种方式 —— 全部构成二十四个单词,依此类推.[10]

8.2.1 亚伯拉罕·伊本艾兹拉的著作

《天地万物》的作者虽曾提到如何计算每次抽取两个字母的组合数,但对组合进行过较详细研究的是犹太教教士亚伯拉罕·伊本艾兹拉(Abraham Ibn Ezra,1090—1167),他是西班牙的犹太裔哲学家,星占学家,和《圣经》的注释者.在一本星占学课本中,伊本艾兹拉讨论了七个"行星"(含太阳和月亮)的可能合数,他相信行星的合对人类生活影响极大.因此,伊本艾兹拉对每一个从2到7的整数 k,算出了组合数 C_k^7,并指出总数目是120.他从最简单的情况开始,即两个一组的合数是21.这个数等于从1到6的所有整数之和,并能用伊本艾兹拉的计算从1到一个具体数字的所有整数之和的法则算出:用这个具体数字之半及二分之一分别乘这个数字后,再相加.用现代术语,就是:

$$C_2^n = \sum_{i=1}^{n-1} i = (n-1) \cdot \frac{n-1}{2} + (n-1) \cdot \frac{1}{2} = \frac{n(n-1)}{2}.$$

为了计算三个一组的组合数,伊本艾兹拉解释道,"我们先将土星和木星以及其它星体之一置于一起,其它星体的数量是5;用5分别乘以它的一半及二分之一,并相加.结果是15.这就是涉及到木星的合数."[11] 亦即,有5个三元组合涉及木星和土星,4个涉及木星和火星而不包含土星,等等.因此,共有 $C_2^6 = 5 \cdot \frac{5}{2} + 5 \cdot \frac{1}{2} = 15$ 种三元组合涉及木星.类似地,为了计算涉及土星而不含木星

的三元组合数,伊本艾兹拉需要算出从剩余的5个中选取2个的组合数 $C_2^5 = 10$. 他然后得出涉及火星而不含木星和土星的三元组合数,并最终得出结论

$$C_3^7 = C_2^6 + C_2^5 + C_2^4 + C_2^3 + C_2^2 = 15 + 10 + 6 + 3 + 1 = 35.$$

伊本艾兹拉接着用类似的方法计算四元组合数. 涉及木星的组合数要求从剩余的6个中选取3个. 涉及土星且不含木星的组合数要求从剩余的5个选取3个. 如此地,最后便有

$$C_4^7 = C_3^6 + C_3^5 + C_3^4 + C_3^3 = 20 + 10 + 4 + 1 = 35.$$

然后伊本艾兹拉直接给出了关于5个,6个,和7个星体的组合数结果. 实质上,他已经论证了 $n = 7$ 的情况,由此很容易概括出一般的组合学法则:

$$C_k^n = \sum_{i=k-1}^{n-1} C_{k-1}^i.$$

注意,伊本慕尼姆(ibn Mun'im)在数十年之后曾对同一结果给出过类似的论述.

除了他的星占学著作,伊本艾兹拉还写过一本算术著作(1146),其中,他介绍了十进制位值系统. 他使用希伯来字母的前九个来表示 1,2,…,9,然后,他便对读者讲述位值制的意义,零的使用(他用圆圈表示零),以及印度 – 阿拉伯系统中的各种计算法则.

8.2.2 莱维本热尔松和归纳法

14世纪早期,莱维·本·热尔松(Levi ben Gerson)在其代表作《计算技术》(Maasei Hoshev, 1321)中,对各种组合学公式给出了细心、严密的证明. 莱维的著作分为两部分,首先是理论部分,其中每个定理都有详细的证明,然后是"应用" 部分,其中给出了完成各种类型计算的明确指令. (莱维在这一部分使用了伊本艾兹拉的"希伯来的" 位值制系统). 莱维在理论部分的开头,有一段关于理论的重要性的非常现代的辩护分析:

> 由于实践工作的真正完善不仅在于知道如何实际完成该项工作,而且在于对它的解释——为什么该项工作要以某种特殊方法来做,同时由于计算技术就是一项实践性工作,那就很明显,重视其理论是恰当的. 还有另外一种原因需要了解这个领域的理论,就是说,这个领域包含多种不同类型的运算,而每种运算本身又涉及许多不同类型的材料,人们很难相信它们全都属于同一学科. 因此,如果不懂理论,就难以领会计算技术,而具备了理论知识之后,完全的精通则很容易. 通晓理论的人们,就能理解如何把理论应用于基于同一基础的各种个例之中. 而对理论无知的人们,就必须孤立地去学习每一种计算,即使二种计算完全一样.[12]

人物小传

莱维·本·热尔松(1288—1344)(Levi Ben Gerson)

莱维可能生于法国南部塞兹河畔巴尼奥尔的乡村,早年生活在附近的奥兰奇镇. 他不仅是数学家,还是天文学家,哲学家,以及《圣经》的注释者. 对其生平所知甚少,仅知他曾与许多重要的耶稣会士保有联系,并应邀创编过一套天文表. 他的各种著作显示出他精通希腊哲学,天文,数学,以及伊斯兰数学传统中的重要部分. 他在天文上的最著名的贡献是他发明了雅各标尺(Jacob Staff),长期被用于测量天体之间的角间距. 它在16世纪的航海家中特别流行(图 8.9).



图 8.9 雅各标尺,由莱维·本·热尔松发明.

当然,就像任何数学著作一样,读者应该具备一些预备知识,譬如欧几里得《原本》的卷 VII,卷 VIII,和卷 IX,“因为我们不打算在本书中重复欧几里得的原话.”但是,莱维坚持对所有他的结果都精心地给出欧几里得式的证明.莱维著作的最重要方面是若干组合学定理.正是在这里,他比伊斯兰前辈们更明确地使用了数学归纳法的要点,他称之为“逐步地无限递进.”一般地,当莱维应用这种证法时,他首先证明归纳步骤,该步骤能从关于 k 的命题推出关于 $k+1$ 的命题,然后指出这个过程从 k 的某一较小的值开始,最后陈述完整的结果.他从未描述过现代的数学归纳法原理,但可以看出他知道怎样使用它.事实上,在该书前面的两个命题中他本质上使用了数学归纳法,这两个命题是论述乘法的结合性和交换性的.

命题 9　两个数的乘积乘以第三个数,其结果等于这三个数中任意两个的乘积乘以另一个数.

命题 10　三个数的乘积乘以第四个数,其结果等于这四个数中任意三个的乘积乘以另一个数.

用现代符号,第一个结果为 $a(bc)=b(ac)=c(ab)$,而第二个将这个结果扩展到四个因子.命题9的证明包含对各种可能的乘积因子的倍数的计算.在命题10的证明中,莱维指出,$a(bcd)$ 含有 bcd 的 a 倍.由命题9,bcd 可被看作 $b(cd)$,因此,乘积 $a(bcd)$ 含有 acd 的 b 倍,或者说,$a(bcd)=b(acd)$.莱维然后将这两个结论推广到任意多个因子的一般情况:“通过逐步递进的过程,就可证明,四个数的乘积乘以第五个数,其结果等于这五个数中任意四个的乘积乘以另一个数.因而,任意个数的乘积乘以另一个数,其结果等于这些数中的任意一个与其余数的乘积的积.”[13] 我们在此便可看到数学归纳法原理的实质.莱维还将这个原理用于证明 $(abc)d=(ab)(cd)$,并总结道,人们可以用同一证明对任意个数的乘积推出下列结论:任意的数都等于它的某两个因子的乘积与其余因子的乘积的积.

莱维并非一贯坚持应用归纳法原理.在该书的中间部分含有许多定理论述各种整数序列的和,这些定理能用归纳法证明.但对其中大多数,莱维用的是其它方法.例如,在证明从1到 n 的整数的和等于$\frac{1}{2}n(n+1)$时(其中 n 是偶数),他利用了第一个与最后一个整数之和,第二个与倒数第二个整数之和,等等均等于 $n+1$ 的思想.当 n 是奇数时,他注意到那些相同的和等于中间数的两倍,便得到同一结果.然而,在证明从1到 n 的整数的立方和公式时,他使用的是归纳法,所用方法使人联想到凯拉吉(al-Karajī)对相同结果的证明.基本的归纳步骤是:

命题 41　从1到某一给定的自然数的和的平方等于该给定数的立方加上从1到比该给定数少一的数的和的平方.[用现代符号,本定理就是$(1+2+\cdots+n)^2=n^3+(1+2+\cdots+(n-1))^2$.]

我们用现代符号来表述莱维的证明.首先,$n^3=n\cdot n^2$.而且,$n^2=(1+2+\cdots+n)+(1+2+\cdots+(n-1))$.(这个结论是莱维的命题30.)那么,

$$\begin{aligned}n^3&=n[(1+2+\cdots+n)+(1+2+\cdots+(n-1))]\\&=n^2+n[2(1+2+\cdots+(n-1))].\end{aligned}$$

但是

$$(1+2+\cdots+n)^2=n^2+2n(1+2+\cdots+(n-1))+(1+2+\cdots+(n-1))^2,$$

故 $n^3+(1+2+\cdots+(n-1))^2=(1+2+\cdots+n)^2$.

莱维接着指出,尽管1之前没有其它数,“它的立方仍是直到它的自然数的和的平方.”换句话说,他给出了用归纳法证明下述结论的第一步.

命题 42　从1到某一给定的自然数的和的平方等于从1到该给定数的立方的和.

莱维的证明完全不是我们所期望的那种归纳法证明.并非从 n 推 $n+1$,而是像凯拉吉一样,他从 n 推 $n-1$.他指出,首先,$(1+2+\cdots+n)^2=n^3+(1+2+\cdots+(n-1))^2$.其最后的被加项也

可由命题 41 得知,等于$(n-1)^3+(1+2+\cdots+(n-2))^2$.照此进行下去,莱维最终得到$1^2=1^3$.结论得证.需要进一步说明的是,虽然命题的叙述用到任意的自然数,但在证明中,莱维仅写出五个数的和,并非我们这里所改用的 n.五个数是用希伯来文的前五个字母表示的.像他的许多前辈一样,莱维无法写出任意多个整数的和,因而是采用可一般化的举例的办法.然而,在莱维的论证中,用归纳法证明的思想是明显的.

《计算技术》的理论部分的最后一段是关于排列和组合,其中的归纳证明也是明显的.该段中莱维指出的第一个结论是,给定的 n 个元素的排列数等于(我们所称的)$n!$:

命题 63 如果若干个不同元素的排列数等于某一给定的数,那么再添加一个元素后的一组不同元素的排列数等于原来的排列数与后来的元素个数的乘积.

本命题用符号表示即 $P_{n+1}=(n+1)P_n$(其中 P_k 表示 k 个元素的全排列数).这个结论为证明命题 $P_n=n!$提供了归纳步骤,虽然莱维直到最后才提到这一结果.他对命题 63 的证明很详细.给定原来的 n 个元素的一个排列,譬如说 $abcde$ 以及一个新元素f,他指出,$fabcde$ 是新集合的一个排列.由于原集合的如此排列有 P_n 个,那么,新集合以f打头的排列也应有P_n个.另外,如果某一原来的元素,譬如 e,被新元素 f 置换,那么集合 a,b,c,d,f 的排列数是P_n,因此,新集合以 e 打头的排列数也应是P_n.因为原集合的 n 个元素的任一个,以及新元素,都能被置于打头位置,由此便知,新集合的所有排列数是$(n+1)P_n$.莱维指出上述$(n+1)P_n$个排列互不相同来结束命题63的证明.然后,莱维断定,"因此,可以证明所给定的一组元素的排列数等于从 1 到所给定的元素的个数的自然数的连乘积.由于两个元素的排列数是 2,且其等于 $1\cdot2$,因此三个元素的排列数是 $3\cdot2$,且其等于 $1\cdot2\cdot3$,照此不断地进行下去,便可得出这个结论."[14] 这就是说,莱维提到了归纳法的第一步,且指出了利用已证明的归纳步骤就可证明整个结论.

证明之后,利用一个计算结果,即 $\mathrm{P}_2^n=n(n-1)$(其中 P_k^n 表示从n 个元素中每次取出k 个元素的排列数),通过关于 k 的归纳法,莱维证明了 $\mathrm{P}_k^n=n(n-1)(n-2)\cdots(n-k+1)$.同前面一样,他将归纳步骤作为定理:

命题 65 如果所有元素的个数已定,且每次从中选取一部分,所选元素的个数也已定,这样的排列数是第三个数,那么,每次再从中多选取一个元素的排列数等于上述第三个数与第一个数和第二个数的差的乘积.

莱维的复杂措词可用现代符号简单表示:$\mathrm{P}_{j+1}^n=(n-j)\mathrm{P}_j^n$.莱维的证明非常类似于命题 63 的证明.最后,他给出了最终结果:"因而已经证明,从所给的元素中选取若干个的排列数等于一列按自然顺序排列的整数的连乘积,其整数的个数等于所选元素的个数,且最后一个整数是所给元素的个数."[15] 为了验证这个结论,莱维首先引用他前面在 $n=7$ 时的结果作为归纳法的第一步,即:从 7 个中选取 2 个的排列数等于 $6\cdot7$.然后,选取 3 个的排列数等于 $5\cdot6\cdot7$(因为 $5=7-2$).类似地,选取 4 个的排列数等于 $4\cdot5\cdot6\cdot7$,"照此办法,可以对任意的数证明这个结论."

在《计算技术》的理论部分的最后三个命题中,莱维完成了他对排列和组合公式的发展.命题 66 指出 $\mathrm{P}_k^n=\mathrm{C}_k^n\mathrm{P}_k^k$,而命题 67 是其简单的改写 $\mathrm{C}_k^n=\mathrm{P}_k^n/\mathrm{P}_k^k$.因为他已经给出了这个商的分子和分母的公式,因此,莱维已经推出了 C_k^n 的标准公式:

$$\mathrm{C}_k^n=\frac{n(n-1)\cdots(n-k+1)}{1\cdot2\cdots k}$$

最后,命题 68 推出 $\mathrm{C}_k^n=\mathrm{C}_{n-k}^n$.

所以,到 1321 年,组合学的基本结果已在欧洲得到.但奇怪的是,在随后的两个世纪中,实际上无人重视组合数学,当组合学后来再度出现时,没有人提到过莱维的贡献(补遗 8.2).

补遗 8.2

有人读过莱维·本·热尔松的著作吗?

虽然《计算技术》是欧洲最早详细论述组合学公式及使用数学归纳法证题的著作,但它在随后几个世纪中似乎未产生任何影响.据目前所知,在后来的欧洲数学著作中均未提及该书,事实上,在随后200年的整个欧洲连组合学公式本身也无人问津.同样,直到17世纪中期帕斯卡的著作出现之前归纳证明也鲜为人知.那么,莱维的书发生了什么意外?曾有人读过它吗?

对后一问题的回答是肯定的.现有十几本该书的手抄本遍布欧洲各地图书馆,还有一本在纽约,大多数传抄于15和16世纪.对于一本中世纪的手稿,这绝不是微不足道的数目.对有些抄本,其抄写者或原始拥有者的姓名是清楚的.事实上,伦敦的抄本曾被15世纪的犹太科学家M·芬齐(Mordecai Finzi)所保存,他曾把艾布卡米勒的著作译成希伯来文.因而最重要的问题是,是否有人读过《计算技术》并用它继续研究过组合学理论吗?

就芬齐来说,没有证据能表明他写过组合学的著述.另一方面,我们确实知道梅森(Marin Mersenne)于1630年代中期在他关于乐理的著作中论述过组合学.而且他的方法论与莱维的著作具有某些相似性.他曾经阅读过莱维的著作或者通过他的许多通信者之一了解过它吗?如果这样,巴黎就应该有人有此手稿的抄本,并且他们应该既懂希伯来文又懂数学.事实上,所有条件都具备.《计算技术》的一个抄本大约在1620年时曾由德桑西(Achille Harlay de Sancy)带到巴黎,此人是法国驻君士坦丁堡的大使.德桑西将该手稿作为他所收集的大量希伯来文著作的一部分捐赠给了天主教的一个崇尚通俗说教的神父团体的图书馆,他本人也加入了该宗教团体的巴黎会所.直到法国大革命期间这个宗教会所于1790年代被查封为止,该手稿一直存放在那里.梅森肯定和该宗教团体中的许多神父有交往,我们也知道其中有些神父懂得希伯来文和数学.但整个线索到此中止.手稿本身没有这方面的记载,也没有该图书馆的记录来说明谁曾借阅过这本书.因此,上述所提问题的答案也许将永不可知了.

8.3　中世纪的代数

虽然组合学理论看来是通过犹太人的传播才在欧洲得以发展,但中世纪欧洲的代数方面的作者则是伊斯兰著作的直接继承者.

8.3.1　比萨的莱昂那多的《算盘书》

欧洲早期的代数作者之一是比萨的莱昂那多,因其代表作《算盘书》(the Liber abbaci)而著称.(算盘并非指计算工具,而是指一般的计算.)该书的第一版出版于1202年,稍微修改后再版于1228年.大量幸存下来的手稿证实了该书的广泛影响.《算盘书》的资料主要来源于莱昂那多曾多次访问过的伊斯兰世界,但他用自己的才华扩充和编排了所收集的材料.该书不仅包含新的印度-阿拉伯数系的计算法则,而且包含各种实用课题中的大量题目,例如利润的计算,货币兑换,和测量.用现行代数课本的标准说法来讲,该书对下列论题作了增补,诸如混合比问题,运动问题,容器问题,中国剩余问题,以及最后部分中,可用二次方程来解的各种形式问题.在这些问题之间分散着一些理论,譬如级数求和的方法和二次公式的几何验证.

莱昂那多在解决问题时使用了大量的方法.事实上,他经常针对特殊问题采用特殊步骤,而非

给出更一般的方法.所经常使用的一种基本方法是古埃及人的"试位法",即先假定一个方便但不精确的结果,然后适当调整以得到正确的结果.莱昂那多也使用了花拉子米的方法解二次方程.其中许多问题都有可能引证出出处.他经常逐字不变地引用花拉子米,艾布卡米勒和凯拉吉这些伊斯兰数学家的问题,其中多数是在他的旅行中从阿拉伯手稿中找到的.有些问题似乎最终源于中国或印度,但莱昂那多可能是从阿拉伯译本了解到的.然而,绝大多数问题属于他本人的设计,并显示出他的创造能力.莱昂那多的某些问题和解法则能体现出这本最有影响的数学著作的风格.

莱昂那多在著作一开始便介绍印度 - 阿拉伯数系:"印度人的九个数字是9,8,7,6,5,4,3,2,1.正如随后所展示的那样,用这九个数字以及符号0,可以表示每一个数."[16] 并确切地给出位置值系统中整数的各数位的名称.莱昂那多然后便论述整数和普通分数的四则运算法则.他的带分数记法不同于现代形式,他是先写分数部分,但他的算法与我们现在的算法很接近.例如,要算83除以$5\frac{2}{3}$,莱昂那多用5乘以3并加上2,得17.他然后用83乘以3,得249,最后用249除以17,得$14\frac{11}{17}$.要算$\frac{1}{5}+\frac{3}{4}$加$\frac{1}{10}+\frac{2}{9}$,莱昂那多把前两个分母5和4相乘,得20,然后再乘以分母9得180.乘以10是不必要的,因为10已是180的因数.然后$\frac{1}{5}+\frac{3}{4}$的180倍是171,而$\frac{1}{10}+\frac{2}{9}$的180倍是58.这两个数的和是229,然后除以180便得最终结果$1\frac{49}{180}$.莱昂那多把答案写为$\frac{1}{2}\,\frac{6}{9}\,\frac{2}{10}1$,用其表示$1+\frac{1}{2\cdot 9\cdot 10}+\frac{6}{9\cdot 10}+\frac{2}{10}$.这种记法也许源于比萨的货币系统.因为1古罗马磅分为20罗马金币,1金币分为12罗马银币,所以,对他来说就很方便,比如把17磅11金币5银币写为$\frac{5}{12}\,\frac{11}{20}17$.除了记法,莱昂那多能够使用他的步骤有效地向读者显示怎样完成复杂的运算,而这些计算可应用于当时在地中海流域流通的多种货币兑换中.

莱昂那多对古典的买鸟问题给出了几种描述.在第一种形式中,他问:如果每只鹧鸪值3个硬币,每只鸽子值2个硬币,两只麻雀值1个硬币,那么,怎样用30个硬币买30只鸟?他开始便说,买4只麻雀和1只鹧鸪的话,便用5个硬币买了5只鸟.同样,买2只麻雀和1只鸽子的话,便用3个硬币买了3只鸟.通过给第一笔交易乘以3和第二笔交易乘以5,他得到能用15个硬币买12只麻雀和3只鹧鸪,也能用15个硬币买10只麻雀和5只鸽子.将这两笔交易相加便给出了所求的答案:22只麻雀,5只鸽子,3只鹧鸪.

另一个是关于坑中之狮的古典问题:坑深50英尺.狮子每天向上爬$\frac{1}{7}$英尺而每晚又下落$\frac{1}{9}$英尺.那么它多少天才能爬出坑?莱昂那多在这里采用了"试位法".他假设答案为63天,因为63可被7和9除尽.那么在63天中狮子要上爬9英尺而掉下7英尺,即实际爬了2英尺.然后,利用比例性,要爬50英尺的话,狮子便需要1575天.(顺便提一下,莱昂那多的答案并不正确.到1571天的最后,狮子离顶端仅有$\frac{8}{63}$英尺.因此,在下一天它将到达顶端.)

书中后面章节的问题更明确地趋向于代数化.例如,假定两个人各有一些钱.第一个人对第二个说,"如果你给我一银币,咱俩的钱数便相同."第二个人对第一个说,"如果你给我一银币,我的钱将是你的十倍."问两人各有多少钱?用现代符号,若x和y分别代表第一个和第二个人的钱,这个问题可化为方程组$x+1=y-1, y+1=10(x-1)$.然而,莱昂那多考察问题的方法有些不同,他引入新的未知数$z=x+y$(总钱数).然后$x+1=\frac{1}{2}z, y+1=\frac{10}{11}z$.两个方程相加得$z+2=$

$\frac{31}{22}z$,由此便知 $z=\frac{44}{9}$,$x=1\frac{4}{9}$,$y=3\frac{4}{9}$.

莱昂那多也轻松自在地论述多于两个未知数的确定问题和不定问题.例如,假定有四个人,其中第一,第二,第三个人共有27银币;第二,第三,第四个人共有31银币;第三,第四,第一个人共有34银币;而第四,第一,第二个人共有37银币.要确定每人有多少银币,就要解由四个方程组成的四元方程组.莱昂那多直接将四个方程相加,得到总钱数的三倍是129银币.每人的钱数然后就很容易算出.另外,在一个类似的问题中,该问题可简化为四个方程 $x+y=27$,$y+z=31$,$z+w=34$,$x+w=37$,莱昂那多首先指出,这个方程组无解,因为用两种不同方式计算总钱数会得到两个不同的答案,61和68.然而,如果把第四个方程改为 $x+w=30$,就可以任意地选择 $x(x\leqslant 27)$ 并依次利用第一,第二,和第三个方程算出 y,z 和 w.

《算盘书》中最著名的问题,兔子问题,被不引人注意地塞在一个完全数问题和上面刚讨论过的问题之间:"由一对兔子开始,一年之后可繁殖成多少对兔子?某人在某地养着一对兔子,并且四周围着墙.假设它们每月生出另一对,而且每对新生的兔子从第二个月起,也开始繁殖,我们想知道,从那对兔子开始,一年之后可繁殖成多少对兔子."[17] 莱昂那多着手计算:第一个月之后将有两对,第二个月之后将有三对.在第三个月,有两对将繁殖,因此,第三个月之后将有五对.在第四个月,有三对将繁殖,因此将有八对.照此进行下去,他指出到第十二个月底将有377对.把数列1,2,3,5,8,13,21,34,55,89,144,233,377列举在页边后,他说每一个数都可由前面两数之和而得出,"因此,便可算出第任意个月后的兔子对数."这个数列现在就叫做斐波那契数列.它具有许多为莱昂那多始料未及的奇特性质,并非仅是与古希腊黄金分割问题的联系.

在最后一章中,莱昂那多在解最终可化为二次方程的问题时,表现出他对伊斯兰先辈的代数学的驾轻就熟.他先像花拉子米一样,依次讨论六种基本类型的二次方程,然后分别对三种混合情况给出方程求解过程的几何证明.证明之后大约有50页的例题,其中大多数来源于花拉子米和艾布卡米勒的著作,包括那些以"把10分成两部分"开头的熟悉的例子.

该书的内容并未反映当时的伊斯兰数学著作中的新进展.事实上,就代数学来说,莱昂那多仅呈现了10世纪的伊斯兰数学,而忽略了11和12世纪的进展.该书的主要价值是,在欧洲提供了对伊斯兰数学最早的全面介绍.它向读者提供了各种各样解决数学问题的方法,这些方法成为后来进一步发展的出发点.

8.3.2 《平方数书》

莱昂那多的另一本较简短的著作《平方数书》(Liber quadratorum,1225)比《算盘书》更加理论化.该书是关于数论的,其中讨论了涉及平方数的各种方程在有理数中的解.该书起因于巴勒莫的约翰爵爷向莱昂那多提出的一个问题,约翰是神圣罗马皇帝腓特烈二世的一位科学随从.约翰提出的问题是,"求一个平方数,无论加上或者减去5,其结果仍是平方数.这个问题的解我已求出,不仅这个问题,经过思考之后,我认为这个解本身以及许多其它解都应是平方数或者平方数之比."[18] 最初的问题,即求 x,y,z 使 $x^2+5=y^2$ 且 $x^2-5=z^2$,作为该书24个命题的第17个已被解决,但莱昂那多首先讲述了平方数以及平方数的和的各种性质.有趣的是,巴勒莫的约翰不仅是数学家,而且是阿拉伯语学者,他一定对凯拉吉著作中的相同问题很熟悉.

为解约翰的问题,莱昂那多引入了他所称的和谐数,即数 n 能表示成 $ab(a+b)(a-b)$(其中 $a+b$ 是偶数)或者 $4ab(a+b)(a-b)$(其中 $a+b$ 是奇数).他指出,和谐数都能被24整除,且仅当 n 是和谐数时,$x^2+n=y^2$ 和 $x^2-n=z^2$ 的整数解存在.因此,最初的问题在整数中无解.然而,由

于 $720 = 12^2 \cdot 5$ 是一个和谐数($a = 5$ 和 $b = 4$),且 $41^2 + 720 = 49^2$ 和 $41^2 - 720 = 31^2$,通过给两个式子都除以 12^2 便知,$x = \frac{41}{12}, y = \frac{49}{12}, z = \frac{31}{12}$ 提供了 $x^2 + 5 = y^2, x^2 - 5 = z^2$ 在有理数中的一个解. 莱昂那多的这套方法不同于凯拉吉的方法(虽然答案一样),但却与同时代的其它伊斯兰著作中的方法类似,例如 10 世纪阿尔 · 哈津(Abū Ja-far al-Khāzin) 的著作.[19]

莱昂那多对他在旅行中所学到的伊斯兰数学显然很精通,并把他的学识传播给欧洲的后继者. 关于数论的《平方数书》,直到几个世纪之后当丢番图的《算术》在欧洲广为传播时,莱昂那多才有了后继者. 另一方面,《算盘书》和《实用几何》中的实用材料被意大利的测地员和计算师们所掌握,这些人在后来的几个世纪中对于把新兴的数学引入意大利有一定的影响. 然而,真正为新数学的产生创造良好的环境则是整整三百年之后的事了.

8.3.3 约丹努斯

约丹努斯(Jordanus de Nemore) 是莱昂那多的同时代人,也曾写过几部重要的数学著作. 虽然我们并不了解他的生平,但相信他曾于 1220 年左右在巴黎教书. 他的著作涉及算术,几何,天文学,力学和代数,似乎他曾以算术理论为基础,主管过一套拉丁文版的四艺丛书的创作工作. 约丹努斯的《算术》与当时在欧洲广为流传的博伊西斯的没有论证分析的算术书大相径庭. 它完全基于欧几里得模型,有一套定义,公理,公设,命题和严格的证明系统. 也和欧几里得一样,约丹努斯没有提供数字上的例子.

《算术》一书共有十卷. 有些论题在欧几里得那里能够找到,如比和比例,素数和合数,欧几里得算法和《原本》卷 II 中的几何代数命题. 也有许多材料不属于欧几里得,包括垛积数和对各种所列举的比率的详细讨论,这些应归于尼科马科斯. 然而,最有趣的是几个古希腊原始资料中所没有的论题. 例如,约丹努斯在卷 VI 中解了一题,实际上该题与莱昂那多《平方数书》中的中心问题相同.

人物小传	约丹努斯(Jordanus de Nemore)
	虽然约丹努斯被认为是中世纪最好的数学家之一,但其生平不详,仅知他在 13 世纪早期与巴黎大学有联系. 几年前,他曾被鉴定为萨克森的约丹努斯(Jordanus de Saxonia),即天主教多明我会的第二任总长老,但近期学术研究说明了这是不可能的.《论给定的数》的翻译者 B · 休斯(Barnabas Hughes) 断定约丹努斯是无父,无母,无亲,无故(Sine patre, sine matre, sine genealogia). 然而,他在一封信中又指出:"[关于为何没有任何生平线索] 我想惟一的解释是,这是一个假名. 那么,会不会约丹努斯其实是一位女性?一位隐姓埋名的希帕蒂娅?13 世纪的女性在作诗,谱曲,和撰写祈祷文方面训练有素. 但是,科学方面又如何呢?"

命题 VI - 12 求三个等差的平方数.

按现代符号,约丹努斯欲求 y^2, x^2, z^2,使 $y^2 - x^2 = x^2 - z^2$. 他的解相当于

$$y = \frac{a^2}{2} + ab - \frac{b^2}{2}, \quad x = \frac{a^2 + b^2}{2}, \quad z = \frac{b^2}{2} + ab - \frac{a^2}{2}.$$

其中 a, b 有相同的奇偶性. 与莱昂那多相比较,约丹努斯仅关心整数解,但没给出任何具体例子. 然而,可以直接看出,按照莱昂那多的定义,约丹努斯定理中的平方数的差就是一个和谐数.

在卷 IX 中,约丹努斯描述的"帕斯卡"三角,在欧洲人的著作中是第一次. 该三角的结构是标准

形式:

> 把 1 置于顶端且第二行是两个 1. 再把第二行的两个 1 相加得 2, 且前后各写一个 1, 使得第三行是 1,2,1. 再把第三行的 1,2 相加得 3, 把 2,1 相加得 3, 并前后各写一个 1, 第四行为 1,3,3,1. 这样, 每一行的数都可用上一行的相应数对来求出.[21]

《算术》的绝大多数中世纪稿本都把此处的三角具体写了出来, 有些甚至呈现到第十行. 约丹努斯然后在命题 IX - 70 中明确地利用这个三角, 来构造已知其比率的数列. 例如, 如果 $a = b = c = d = 1$, 那么, 数 $e = 1a = 1$, $f = 1a + 1b = 2$, $g = 1a + 2b + 1c = 4$, 和 $h = 1a + 3b + 3c + 1d = 8$ 就形成了一个具有固定比率 2 的连比例. 类似地, $k = 1e = 1$, $l = 1e + 1f = 3$, $m = 1e + 2f + 1g = 9$, 和 $n = 1e + 3f + 3g + 1h = 27$ 形成一个具有比率 3 的连比例.

至少从现在还保存的稿本的数量可看出, 约丹努斯的《算术》流传很广. 类似地, 他的主要代数著作《论给定的数》(De numeris datis) 在中世纪欧洲也拥有广泛的读者.《论给定的数》是一部代数方面的分析著作, 虽然以 13 世纪早期之前已传入欧洲的伊斯兰代数知识为基础, 但却具有不同的思想风格. 它似乎以欧几里得的著作为模式, 而约丹努斯当时完全可以见到克雷莫纳的杰拉德所译的拉丁文版, 其中就有这样的问题, 如果给定某些量, 那么另一些量也能被确定. 然而,《论给定的数》中的问题是代数的而非几何的形式. 约丹努斯的证明也是代数的, 或算术的. 事实上, 他的目的之一明显是想把新代数建立在算术, 而非几何的基础上. 他也以逻辑方式组建其著作, 但与欧几里得模式, 甚至他自己的《算术》的一个主要区别是, 对大多数理论结果都提供了数字例子.

虽然许多实际问题和数字例子来源于伊斯兰代数, 但约丹努斯使它们适合于自己的用途. 特别地, 他做出用字母表示任意数的重大改变. 约丹努斯的代数不再完全是文字的. 这并不是说他的符号表示是现代形式. 他按字母顺序选取字母, 而不区分用哪些字母表示已知量和未知量, 并且没有用符号表示运算. 有时用两个字母表示一个数. 有时却用字母对 ab 表示两个数 a 和 b 的和. 算术基本运算全用文字表示. 而且, 约丹努斯没有使用新的印度 - 阿拉伯数字. 所有数都写成罗马数字的形式. 然而, 对于代数技巧主要进展作用重大的符号表示思想, 至少其雏形, 已在约丹努斯的著作中出现了.

《论给定的数》共有四卷, 我们从它的 100 多个命题中举出几例, 来了解约丹努斯的贡献. 同欧几里得一样, 约丹努斯以标准形式给出每一命题. 一般的理论阐释之后, 再用字母方式给予重述. 通过一般的法则, 把代表数的字母处理为规范的形式, 而用这种规范的形式能够容易地找到一般的解. 最后, 用抽象解法的一般步骤对具体的数字例子做出计算.

命题 I - 1　如果某一给定的数被分成两部分, 且其差是已知的, 那么, 每一部分能够被确定.

约丹努斯的证明很直接:"这就是说, 较小的部分和其差构成较大的. 因此, 较小的部分与它本身和其差构成全部. 所以, 从全部中减去其差便剩较小部分的两倍. 除以 2, 较小的部分便被确定; 当然, 较大的也被确定. 例如, 把 10 分成两部分, 其差为 2. 10 减去 2 剩 8, 其半是 4, 这就是较小的部分. 另一部分便是 6."[22]

用现代符号表示, 约丹努斯的解相当于两个方程 $x + y = a$, $x - y = b$ 的解. 约丹努斯首先指出, $y + b = x$, 因此 $2y + b = a$, 所以 $2y = a - b$. 故 $y = \frac{1}{2}(a - b)$ 且 $x = a - y$.

约丹努斯在卷 I 的许多问题中都使用了这个最早的命题. 例如

命题 I - 3　如果某一给定的数被分成两部分, 且其积是已知的, 那么, 每一部分必能被确定.

这个命题是巴比伦人的标准问题之一: $x + y = m$, $xy = n$. 然而, 约丹努斯的解法不同于巴比伦人的古典方法, 他使用的是下面所指出的符号表示法: 假设给定的数 abc 被分成 ab 和 c. 假设 ab

乘以 c 是 d,且 abc 自乘是 e.令 f 是 d 的四倍,且 g 是 e 和 f 的差.那么,g 是 ab 和 c 的差的平方.所以它的平方根 b 是 ab 和 c 的差.现在,由于 ab 和 c 的差及和都已给定,按照第一个命题,ab 和 c 就能被确定.约丹努斯的数字例子为,两部分之和是10且其积是21.他指出,84是21的四倍,100是10的平方,且16是它们的差.那么,16的平方根4便是10的两部分之差.用第一个命题的证明过程,10减去4得6.所以,3是所求的较小部分,而7是较大的.

约丹努斯的解法翻译成现代的符号表示,相当于使用恒等式 $(x-y)^2=(x+y)^2-4xy=m^2-4n$ 确定 $x-y$,且把问题化为命题 I-1.那么,解为 $y=\frac{1}{2}\left(m-\sqrt{m^2-4n}\right)$,$x=m-y$.约丹努斯的方法似乎对他来说是新创,因此他在全书反复地使用他自己的方法.他甚至在解卷 I 中相当于二次方程的问题时,也使用了不同于伊斯兰代数学家的标准方法.然而,数字例子本身却看起来很相似.事实上,卷 I 中每个命题都是关于把一个数分成两部分,且每个例子中惟一被分的数是10.解法在某种程度上不同于伊斯兰课本中的方法,但题目很清楚地依赖于花拉子米的问题.

其余三卷中的许多命题论述给定比例的数.这些命题显示出约丹努斯能流畅地处理欧几里得《原本》卷 V 和卷 VII 中的比例法则,其中大多数也出现在他自己的《算术》中.

命题 II-18 无论某一给定的数被分成几部分,只要已知其连比例,那么,每一部分就能被确定.

因为约丹努斯和他的同时代人一样,无法表示任意多的"部分",因此在证明中,他论述的是一个被分成三部分的数:$a=x+y+z$.且 $x:y=b$ 和 $y:z=c$ 都是已知的比率.约丹努斯指出,比率 $x:z$ 也是已知的.由此便知 x 比 $y+z$ 的比率,因此 a 比 x 的比率也可知.因为 a 是已知的,所以 x,然后 y 和 z 就能被确定.他的例子能使我们理解他的文字描述.数60分成三部分,其中第一部分是第二部分的2倍,第二部分是第三部分的3倍.即,$x+y+z=60$,$x=2y$,$y=3z$.那么 $x=6z$,由此便知 $y+z=\frac{2}{3}x$.因此 $60=1\frac{2}{3}x$,且 $x=36$,$y=18$,$z=6$.可以看出,约丹努斯在需要时能方便地转化比率关系,而且知道怎样合并它们.

卷 IV 中有三个命题给出了二次方程的三种标准形式,它们都以代数证明,而非几何验证的形式出现.然而对于这些问题,约丹努斯用的是伊斯兰的标准算法,但却以他自己的符号来表述.

命题 IV-9 如果一个数的平方加上某一给定的数等于该数与另一定数的乘积,那么,该数可能取两个值.

这就是说,约丹努斯断言方程 $x^2+c=bx$ 可能有两个解.他然后给出解方程的过程:取 b 的一半并平方之,得 f,且令 g 是 x 与 $\frac{1}{2}b$ 的差,即,$g=\pm\left(x-\frac{1}{2}b\right)$.那么 $x^2+f=x^2+c+g^2$ 且 $f=c+g^2$.约丹努斯最后断定,可以用 $\frac{b}{2}$ 减去 g,或用 $\frac{b}{2}$ 加上 g 得出 x.他的例子使他的符号过程更加清楚.欲解 $x^2+8=6x$,他平方6的一半得9,然后从中减去8剩1.1的平方根仍是1,这便是 x 与3的差.所以 x 是2或4.

卷 IV 中有两个二次方程问题实质上是方程组 $xy=a$,$x^2+y^2=b$ 和 $xy=a$,$x^2-y^2=b$.和该书前面的情形一样,此处所给例子推出的也是正整数答案.虽然约丹努斯在解题过程中经常使用分数,但他总是细心组织材料,以便最后的答案为整数.事实上,艾布卡米勒的《代数》当时已有拉丁文版,如果他读过的话,约丹努斯应该见过这种问题的非整数,甚至非有理数的解.然而,在构造他的例子时,他排除了这样的解.这足以说明,约丹努斯仍受欧几里得的影响,感到无理"数"不应该属于一本以算术为基础的著作.所以说,虽然在分析法的使用方面,在向一般性的不懈努力方面,

以及在符号化方面,《论给定的数》比伊斯兰著作都有所进步.但该书又返回到古希腊人把数和量严格区分的思想上,而约丹努斯的伊斯兰前辈已经背离了这种思想.这说明,虽然约丹努斯把新的伊斯兰材料传到了欧洲,但他的目的是要给读者提供一种尽可能基于古希腊原则的数学.

8.4　运动的数学

在 13 世纪,约丹努斯的代数工作并未得到进一步发展,虽然到 13 世纪中期曾有一些追随者在巴黎出现.也许欧洲当时还未做好恢复纯数学研究的准备.然而,到 14 世纪早期,当学者们试图澄清亚里士多德物理学论著中的某些论断时,数学的某些其它方面便开始在牛津大学和巴黎大学发展起来(补遗 8.3).

补遗 8.3　中世纪的大学

12 世纪后期,欧洲出现了大学端倪,对后来整个科学及数学科目的发展产生了巨大的影响.我们对这些最早的大学起源问题不可能确定出具体的日期.它们是从中世纪西欧的牧师和学生的同业行会逐步演变而成.这些最早的机构出现在巴黎,牛津,和波伦亚.巴黎大学产生自圣母院的天主教学校.那些教师和学生逐渐把他们自己分成四种系科,即技艺,神学,法律和医学.虽然有证据表明该大学在 12 世纪后期实际已经存在,但最早的官方特许状日期是 1200 年.牛津大学并非源于某一支教会学校,而是源于从巴黎返回的一批英国学生.同样,虽然该大学在 12 世纪后期实际已存在,但最早的官方批准文书是 1214 年.波伦亚大学最早是一所法律学校,该学校成立于 11 世纪.这所意大利大学最初是由一批学生自发组成的协会,由学生选举教师及其它官员.然而,学生的权力逐渐削弱,因为薪水由波伦亚市政当局支付且教师要实施考试.

所有这些大学的技艺类课程都基于古代的三科(逻辑,语法和修辞)和四艺(算术、几何、音乐和天文)(图 8.10).技艺专业的这种研究为学生进一步学习法律,医学,或神学提供了准备.技艺课的核心是逻辑研究,这方面的主要课本是亚里士多德的逻辑著作,而这些著作到当时均已被译成拉丁文.这样做是因为逻辑能教导哲学和科学探索的方法.逐渐地,亚里士多德在其它方面的著作也被用于教学.在长达数世纪期间,这位伟大哲学家的著作是整个技艺类课程的基本焦点.如果其他人的著作能帮助理解这位最多产博学的哲学家,也会被研究.特别地,数学研究也仅仅是因为它与亚里士多德的逻辑和物理著作相关.数学课程本身,即四艺,通常含有:算术,取材于博伊西斯对尼科马科斯著作的改编本或中世纪的某一部算法书;几何,取材于欧几里得和实用几何;音乐,取材于博伊西斯的一部著作;天文,取材于托勒密的《大成》和一些较晚的伊斯兰天文著作的拉丁文译本.





图 8.10　四艺:算术、几何、天文和音乐.

8.4.1 比率的研究

一种新的数学思想起源于对施加于一物体的作用力 F、阻力 R、和速率 V 之间的关系的探究. 中世纪物理学的一个基本假设是 F 必须大于 R,运动才会产生.(中世纪的哲学家并未尝试用任何特定的单位去测量这些量.)蕴涵在亚里士多德原文中的这些量之间的最简单的关系可以这样来表述,F/R 与 V 成正比例. 然而,这个数学关系会立即导致与基本假设的矛盾. 因为,如果 F 保持不变,不断地加倍 R 就等于不断地减半 V. 正速率减半时其正号不变,而对 R 的加倍最终会使 R 大于 F,这样就发生与 F 必须大于 R,运动才会产生的假设的矛盾.

牛津大学莫顿学院的布雷德沃丁(Thomas Bradwardine,1295—1349),在 1328 年的《论运动中速率的比例》(Tractatus de proportionibus velocitatum in motibus) 中,对这个二难推论提出了一种解决方法,即,对亚里士多德原话给出一种"正确" 阐释. 上述法则的含意为,有两个力 F_1, F_2,两个阻力 R_1, R_2,和两个速率 V_1, V_2,满足下面的方程式:

$$\frac{F_2}{R_2} = \frac{V_2}{V_1}\frac{F_1}{R_1}.$$

布雷德沃丁提出它应当被以下用现代符号表示的关系式所替代:

$$\frac{F_2}{R_2} = \left(\frac{F_1}{R_1}\right)^{V_2/V_1}.$$

换句话说,乘积关系应当被指数关系所替换. 这个方法确实消除了上述谬论. 最初给定 $F > R$(或 $F/R > 1$),在这种情况下,把速率减半就等于取比率 F/R 的[平方]根. 因此,F/R 将仍然大于 1,R 也就永远不会大于 F. 然而,无论布雷德沃丁,还是同时代的其他任何人都没有就这个关系式试着给出任何实验性的验证. 莫顿的学者们想对世界作出一种数学解释,而非物理解释. 到下一世纪中期,布雷德沃丁的思想作为一种物理原理终于被抛弃了,但是它背后所隐藏的数学却引出了新的重要思想. 对它们的论述需要对比率进行系统的研究,特别是复合比率(或乘积比率) 的思想.

直到 14 世纪,复比仍被呈现为古典希腊的形式. 因此要处理 $a:b$ 和 $c:d$ 的复比,就应找到一个量 e,使 $c:d = b:e$. 那么所求复比便是 $a:e$. 然而,更明晰的关于比率乘积的解释被逐渐引人. 例如,布雷德沃丁在牛津的同时代人,沃灵福德的理查德,在他的《四分集》的卷 II 中,就定义了比率以及它们的乘法和除法:

1. 比率是两个同类量之间的相互关系.
2. 用两个同类量之一除以另一量,所得的商叫做被除数与除数的比值.
3. 复比的值是各单比的值的连乘积.
4. 两个比率相除,其比值为两已知比率的比值的商.[23]

这里需要作几点重要解释. 首先,理查德强调比率只能用于同类量之间. 这种欧几里得思想意味着速率不能被认为是距离与时间的比率. 其次,比值指比率的结果,这一点源于尼科马科斯(Nicomachus),现已在欧洲通行. 例如,比率 3 : 1 叫做三倍的比率,而 3 : 2 叫做一倍半的比率. 最后,定义 3 和 4 反映出理查德不同于欧几里得,数字的乘法和除法被用于比率的乘法和除法. 虽然他将比率 4 : 16,16 : 2,和 2 : 12 复合后得 4 : 12,但他指出,因三个比率依次为 1 : 4,8 : 1,1 : 6,也可以这样来复合,先用 4 除 8 得 2,再用 6 除 2 得 1 : 3,并作为最终结果. 这样,就可将分数的乘法法则用于"复合" 比率.

尼科尔 · 奥雷姆(Nicole Oresme,1320—1382),一位曾学习和任教于巴黎大学的牧师和数学家,

在他的《比率算法》(Algorismus proportionum)和《论比率之比率》(De proportionibus proportionum)中对比率做过非常详细的研究.除了以传统方法呈现复比之外,奥雷姆确切地指出,也可通过前项乘前项和后项乘后项来复合比率.因此 4 : 3 与 5 : 1 复合是 20 : 3.两个方法之间的联系可能是,$a : b$ 可表示为 $ac : bc$,$c : d$ 可表示为 $bc : bd$,所以,$a : b$ 与 $c : d$ 的复比就是 $ac : bc$ 与 $bc : bd$ 的复比,亦即 $ac : bd$.在给定两个比率相乘的方法之后,奥雷姆还指出,可通过把前项和后项分别相乘,得两个比率的除法.所以,$a : b$ 除以 $c : d$ 的商就是比率 $ad : bc$.

在定义了两个比率的乘积之后,奥雷姆便讨论某一比率的自乘.因此,$a : b$ 自乘 n 次便给出现在所说的 $(a : b)^n$ 的值.更重要地,对于给定的比率,奥雷姆设计了一套文字语言来对应现在所说的该比率的"根".例如,因为 2 : 1 是二倍的比率,奥雷姆便称与其自身相乘等于 2 : 1 的那个比率为二倍比率之半.(按现代术语学,即指 $(2:1)^{1/2}$).类似地,他称 $(3:1)^{3/4}$ 为三倍比率的四分之三.然后,奥雷姆发展了关于这些比率的运算.例如,要计算 $(2:1)^{1/3}$ 乘以 3 : 2,奥雷姆把第二个比率立方,得 27 : 8,再乘 2 : 1,得 27 : 4,然后,把 27 : 4 看作分数取其立方根,便得 $\left(6\frac{3}{4}\right)^{1/3}$.类似地,要计算 $(2:1)^{1/2}$ 除以 4 : 3,他用 2 : 1 除以 4 : 3 的平方根,亦即 16 : 9,得 9 : 8,然后再取其平方根,便得 $(9:8)^{1/2}$.从一定程度上讲,奥雷姆首次引入了分数指数及其运算法则.

奥雷姆甚至想涉及我们所说的无理指数.他感到"比率应该像关于除法的连续量一样",亦即,该比率的任何可能"部分"都应该被得到.因此,"应该有某一比率对应于二倍比率的这样部分,它不是二倍比率之半,也不是三分之一部分,或四分之一部分,或三分之二部分,等等,而是与二倍比率不可公度,进而与二倍比率可公度的任何比率都不可公度".[24] 由于奥雷姆未曾定义无理指数,故只能概述其思想.也就是说,他感到即使 r 不是有理数,$(2:1)^r$ 形式的比率也应该存在."进一步,用相同的推理可知,应该有某一比率既与二倍比率、三倍比率不可公度,又与和这些比率可公度的任何比率不可公度…….应该有某一无理比率,与任何有理比率均不可公度.也就是说,既然有某一比率与两个[有理]比率不可公度,且有某一比率与三个有理比率不可公度,等等,那么,就有某一比率与任何有理比率均不可公度…….然而,我不知道如何展示这个结果."[25] 用现代思想表述,奥雷姆所想明确表示的是,因为实数直线是连续的,而 2 的分数幂没有穷尽所有实数,因此,肯定有 2 的非分数幂等于所没有包含的那些实数.事实上,在该书后面部分,他谈到一个定理,意思是无理比率比有理比率更为普遍:

命题 III－10 由于在所指定的大量未知比率中,绝大多数情况下其中某两个比率均是不可公度的,所以,一般来说,所指定的两个未知比率很可能是不可公度的.

虽然奥雷姆未对此作出正式证明,但他举例说,在从 2 : 1 到 101 : 1 的所有整数比率中,若对其中某两个比率按照指数作比较,并且总是较大的比率比较小的,则共有 4950 种情况,但仅有 25 种情况具有有理指数.例如,$4:1=(2:1)^2$ 且 $8:1=(4:1)^{3/2}$.另一方面,没有有理指数 r,使 $3:1=(2:1)^r$.奥雷姆然后用概率论证得出结论:占星学是不可靠的.他的论据是表示各种天体运动的两个未知比率之间的比为无理数的概率极大.既然没有关于行星会冲的精确循环模式,而占星学却又依赖于这种永无休止的循环,故这门"科学"的整个基础是不可信的.

8.4.2 速率

把亚里士多德关于运动的思想转化为数量结果的努力也导致了新的数学.特别是布雷德沃丁和莫顿学院的另一位学者海特斯伯里(William Heytesbury),在 14 世纪早期发展了这些思想.早期的希腊数学家,包括奥托利科斯(Autolycus)和斯特拉托(Strato),已经涉及到匀速的概念以及一定程

度的加速运动,但从未把速度和加速度作为可以测度的独立的量.他们仅仅通过比较距离与时间来讨论速率,因此,被比较的实质上是某些不同时间段的平均速率.

然而在 14 世纪,作为可测度的客观实在的速率特别是瞬时速率已开始出现.布雷德沃丁在他的《论连续》(Tractatus de continuo)(大约 1330 年)中,把运动的"级(grade)"定义为能够反映运动的"多"和"少".[26] 布雷德沃丁然后说明了如何比较速率:"对于两个连续运动而言,在同一或相等的时间间隔内,速率和[运动]所越过的距离成比例,即,一个运动的速率比另一运动的速率等于该运动所越过的距离比另一运动所越过的距离…….而对所越过的距离一样或相等的两个局部运动而言,速率与时间成反比例,即,第一个速率比第二个速率等于第二个速率的时间比第一个的时间."[27] 换句话说,如果两个物体以匀速 v_1,v_2 运动,时间依次为 t_1,t_2,距离依次为 s_1,s_2,那么(1)若 $t_1 = t_2$,则 $v_1 : v_2 = s_1 : s_2$,(2)若 $s_1 = s_2$,则 $v_1 : v_2 = t_2 : t_1$.这说明布雷德沃丁认为,均匀的速率本身是一种量,可以与其它的速率作比较.

几年之后,海特斯伯里在他的《解答诡辩之法则》(Regule solvendi sophismata,1335)中,对非均匀运动的物体的瞬时速率,给出了一个详细的定义:"在非均匀运动中,对于任意指定的时刻,该时刻的瞬时速率可被描述为,假设该物体以该时刻的速率作匀速运动,在单位时间内所走过的路程."[28] 在给出这个精确定义之后,海特斯伯里举例解释说,即使两个质点,在某一时刻的瞬时速度相同,它们在相等的时间内,所发生的位移也未必相同,因为在其它时刻,它们的速率可能完全不同.

海特斯伯里在同一章节还论述了加速度:"对于任何运动而言,如果在任意相等的时间间隔内,所得到的速率的增量相等,那么,它就是匀加速运动…….如果在某两个相等的时间间隔内,所得到的速率的增量不同,那么,它就是非匀加速运动…….由于匀加速运动的不同时刻的速率不相同……,因此可以将一运动物体从静止状态匀加速运动到任何指定的速率."[29] 这个陈述不仅提供了匀加速的确切定义,而且,至少以最初形式,说明速率随时间而变.换句话说,海特斯伯里把速率描述为时间的"函数".

如何确定匀加速运动物体的位移?其解答,也就是现在所说的平均速度法则,最早也是由海特斯伯里在同一本著作中给出:"当一运动物体从初速为零被匀加速运动到某一给定的速率时,在这段时间内该物体的位移就等于,假设该物体以最后所给定的速率作匀速运动的话,在这段时间内所走过的距离的一半…….这个匀加速运动整体来说,完全相当于以最终速率之半做匀速运动."[30] 用现代符号,如果匀加速运动的物体初速为 0,时间为 t,加速度为 a,那么它的最终速率 $v_f = at$.海特斯伯里所说的是,这个物体的位移为 $S = \frac{1}{2}v_f t$.若把第一个公式代入第二个,便可得到标准的现代公式 $S = \frac{1}{2}at^2$.

海特斯伯里利用对称性推理给出这个平均速度定理的一个证明,作为模型,他假定物体 d 从静止状态在一小时内匀加速运动到最终的速率为 8.(数字 8 并不代表任何具体的速度,仅是为了举例的方便.)然后他考虑另外三个物体,a 在一小时内始终以速度 4 作匀速运动,b 在半小时内从速度 4 匀加速运动到速度 8,c 在半小时内从速度 4 匀减速运动到速度 0.首先,他说,物体 d 在前半个小时的位移和 c 的位移一样,而在后半个小时的位移和 b 的位移一样.因此,d 在整个小时内的位移应等于 b 和 c 在半小时内的位移之和.其次,他推理说,既然 b 所增加的完全等于 c 所减少的,因此它们合并起来在半小时内的位移就完全相当于两者都以速度 4 匀速运动的位移.而后者与 a 在整个小时的位移相同.这说明 d 和 a 在该小时的位移完全一样,在海特斯伯里看来,平均速度定理便得到了证明.然后,他证明了一个简单的推论,即,物体 d 在后半小时的位移是其前半小时位移的三倍.

在莫顿学院的其他同时代学者开始探索用线段表示速率以及其它各种数量的思想.事实上,这种思想来自于亚里士多德,因为这位希腊哲学家在区别两种类型的数量时,把诸如时间,距离,和线段的长度这些概念都视为连续的量.既然都是无限可分的,因此就有理由把有点抽象的速率概念,在它本身已被量化之后,用线段这种具体的几何思想来替代.于是,大小不同的速率可以用不同长度的线段表示.奥雷姆把这种思想发展为,对依据时间而变化的速率引入一种二维的表示.事实上,大约在 1350 年,在他的《论质量与运动的形态幅度》(Tractatus de configurationibus qualitatum et motuum)中,奥雷姆甚至把这种思想推广到其它的情况,其中某一给定的量依据距离或时间而变化.

奥雷姆开篇便对为什么可用线段来表示速率一类的量作出解释:

> 除了[离散的]数外,一切可测度的东西都可视为连续的量.因此,可设想利用点,线,面,或它们的性质来测量这样的量.因为对点、线、面来说,它们的度量或比率,正如亚里士多德已经做到的那样,本来就能确定,而对其它事物来说,我们认为也可以借助于思维将它们与点线面等[几何实体]相联系而确定其度量或比率.虽然不可分的点或线并非客观实在,仍有必要对它们做出数学的假设,以帮助对事物的度量以及对其比率的理解.因此,凡能被连续地获得的每一种量都应该被设想为垂直于某点的直线段.[31]

奥雷姆利用这些直线构建了他所称的"形态幅度",一种由画在底线上的所有垂线组成的几何图形.在速率的情况下,底线代表时间,而垂线表示每个时刻的速率.整个图形表示速率的整体分布,奥雷姆将其解释为,代表运动物体所移动的总距离.奥雷姆没有使用我们所说的坐标.没有用固定的长度表示某一给定大小的速率.重要的思想仅在于:"相等的线段表示相等的速率,二倍的线段表示二倍的速率,依次类推,按比例进行下去."[32]

奥雷姆然后把一个恒定的量,例如,匀速运动的物体,表示为一个矩形,因为在每一点速率是一样的(图 8.11).矩形的面积表示总的位移.一个物体从静止状态出发,然后做匀加速运动,则其总位移是直角三角形的面积(图 8.12).正如奥雷姆所说,"匀加速运动就是,如果[在底线上]取三个点,则第一点和第二点的距离与第二点和第三点的距离之比等于第一点处的速率超出第二点处的速率的增量与第二点超出第三点处的速率的增量之比,并称第一点最大速率点."[33] 这种比率的相等性自然确定一条直线,即直角三角形的斜边.最后,对"非匀加速"的量,例如非匀加速的速率,用"顶线"是曲线而不是直线的图形来表示(图 8.13).亦即,奥雷姆实质上已产生了利用曲线表示速率和时间之间的函数关系的思想.事实上,他说"利用这种抽象出来的图形关系,能更好,更清楚,更容易地理解上述各点速率的变化."[34] 换句话说,对各种量的几何表示提供了研究它们的最佳方式.

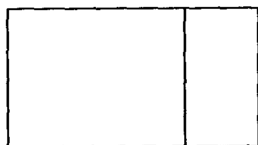


图 8.11　匀速运动.

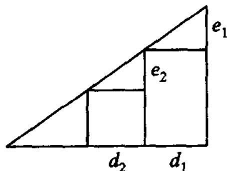


图 8.12　匀加速运动,其中 $d_1 : d_2 = e_1 : e_2$.

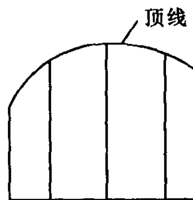


图 8.13　非匀加速运动.

给出物体运动的这种表示法之后,奥雷姆应该能很容易地给出平均速度定理的几何证明.正如

莫顿的学者们所说,如果三角形 ABC 表示初速为零的匀加速运动的物体的形态幅度,且 D 是底线 AB 的中点,那么,垂线 DE 表示中点处的速率,应为最后的速率之半(图 8.14).三角形 ABC 所表示的总位移就等于矩形 $ABGF$ 的面积.

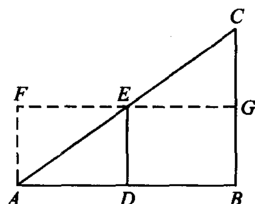


图 8.14　平均速度定理的几何证明,应归功于奥雷姆.

奥雷姆的几何技巧在大约 250 年之后,重现在伽利略(Galileo) 的著作之中.两者的主要区别在于,伽利略思考的是自由落体的物理规律,而奥雷姆是抽象地对匀加速运动作研究.这种抽象性,在奥雷姆分析那些速率可以无限增大的情况时,是非常明显的.例如,他讨论了这样的问题,物体在时间间隔 AB(不妨取为单位 1) 的前一半的速率是 1,在随后的四分之一是 2,在随后的八分之一是 3,在随后的十六分之一是 4,等等,求其总位移.就是说,他在求下列无穷级数的和:

$$\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2+\frac{1}{8}\cdot 3+\cdots+\frac{1}{2^n}\cdot n+\cdots.$$

他的结论是,这个和,或总位移为 2,即"前一半时间所走距离的四倍."[35] 他所给的几何证明是非常漂亮的.令 CD 等于 $AB(=1)$,他作出一个以 CD 为边的正方形,并对其按 2 : 1 的比率进行无限分割(图 8.15).亦即,E 表示正方形的一半,F 为四分之一,G 为八分之一,等等.把矩形 E 置于 AB 上的正方形的上面的右半部,再把 F 置于新形态图上面的右四分之一处,而 G 在上面的八分之一处,等等.则显然,新形态图的总面积,它表示总位移,不仅等于该无穷级数的和,而且等于两个原始正方形的面积的和.

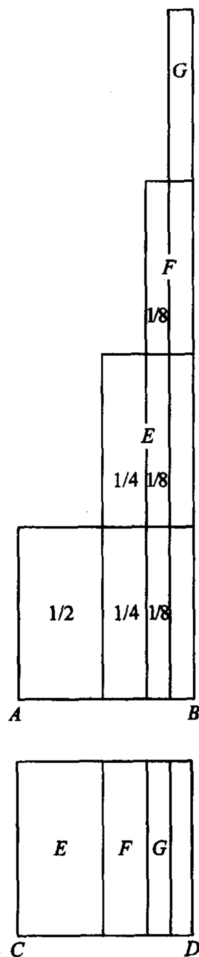


图 8.15　奥雷姆的求和法,$\frac{1}{2}\cdot 1+\frac{1}{4}\cdot 2+\frac{1}{8}\cdot 3+\cdots+\frac{1}{2^n}\cdot n+\cdots$.

奥雷姆用几何方法表示速率以及其它量的思想,在随后一个世纪的其他学者的著作中得以继承.然而,没有人能将奥雷姆关于匀加速量的距离表示推广到更复杂的情形.最后,甚至连这个思想本身也被遗忘了.许多中世纪欧洲的重要数学家的思想也有着类似的命运.他们的著作没有得到研究,他们的新思想在几个世纪后,不得不被重新发现.这种停滞不前在最早的大学以及随后几个世纪所建的大学的数学课程中体现得甚为明显,数学研究仅仅是为了帮助学生理解伟大哲学家的著作.虽然一个像奥雷姆这样的人也许可以将这些思想推向前进,但这样的学者在当时实在是凤毛鳞角.另外,由于黑死病和百年大战在法国和英国也引起了显著的学术倒退.因此,只是在意大利和德国,才有一些数学家对中世纪法国和英国数学家的思想进行研究.文艺复兴时期的新思想正是诞生在那里.

习　题

阿尔昆《敏锐青年之命题》中的问题

1. 三根管子给一木桶共注入水 100 metreta.其中一根管子注入木桶容量的三分之一加 6 modii;另一管子注入容量的三分之一;第三根管子仅注入容量的六分之一.每根管子各注入多少 sextarii?(其中 1 metreta 等于 72 sextarii,1 modius 等于 200 sextarii.)[36]

2. 某人要将一只狼，一只羊和一棵洋白菜运过河. 而小船只能运载这个人和另外一种东西. 羊不能单独和白菜留在一起，也不能把狼和羊单独留下. 这个人应该怎么办？

3. 一只狗在追前方 150 步远的兔子. 如果狗每次跑 10 步而兔跑 6 步，那么狗跑多少步便能追上兔子？

萨瓦苏达的问题

4. 在直径为 $10\frac{1}{2}$ 的圆上，已知某一弦长是 6，求该弦所对应的弧长.

5. 求出由练习 4 中的弦所确定的弓形面积.

6. 在直径为 33 的圆中，某一弧长是 $5\frac{1}{2}$，求该弧所对应的弦长.

7. 如果长度为 8 的弦到圆周的距离是 2，求该圆的直径.

8. 在对角线为 10 的矩形中，如果对角线比某一边多 2，求矩形的长，宽和面积.

9. 如果某一菱形的两条对角线分别是 16 和 12，求其边长是多少？

实用几何问题

10. 说明 $A = \frac{3n^2 - n}{2}$ 给出了求第 n 个五边形数的公式. 计算边长 $n = 1, 2, 3$ 的正五边形面积，并与第 $(n+1)$ 个五边形数的值比较. 所给公式的近似程度如何？

11. 如图 8.6 所示，用两次观测法求塔高. 在点 S_1 测得塔高与距离 d_1 之比为 2:5. 在点 S_2 测得塔高与距离 d_2 之比为 2:7. 如果两观测点的距离是 50 米，求塔高是多少？

12. 应用莱昂那多的弦表求弧长：如果在一直径为 10 的圆中，某一弦长是 8 rods，3 feet，$16\frac{2}{7}$ unciae. 求该弦所对应的弧长.

13. 莱昂那多《实用几何学》中有一题：如图 8.16 所示，已知圆内接四边形的 $ab = ag = 10$ 且 $bg = 12$，求该圆的直径 ad.

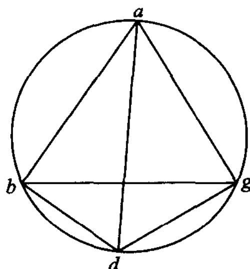


图 8.16　确定圆的直径，选自比萨的莱昂那多的原著.

三角学问题

14. 像沃灵福德的理查德那样，建立一个公式来计算三个弧的和所对应的弦. 并将其转化为求三个弧的和的正弦公式.

15. 试证莱维本热尔松《三角学》中的如下定理：已知任一三角形的三边长，便可推出它的三个角. 先从某一顶点向对边（或对边的延长线）作垂线，并说明怎样计算三个角.

算术问题

16. 用数学归纳法证明组合学公式：

$$C_k^n = \sum_{i=k-1}^{n-1} C_{k-1}^i.$$

17. 证明莱维本热尔松《计算技术》中的命题 30：

$$(1 + 2 + \cdots + n) + (1 + 2 + \cdots + (n-1)) = n^2.$$

18. 证明《计算技术》的命题 32：

$$1 + (1+2) + (1+2+3) + \cdots + (1+2+\cdots+n) = \begin{cases} 1^2 + 3^2 + \cdots + n^2, & n \text{ 是奇数}, \\ 2^2 + 4^2 + \cdots + n^2, & n \text{ 是偶数}. \end{cases}$$

19. 证明《计算技术》的命题 33：

$$(1 + 2 + 3 + \cdots + n) + (2 + 3 + \cdots + n) + (3 + \cdots + n) + \cdots + n = 1^2 + 2^2 + \cdots + n^2.$$

20、证明《计算技术》的命题 34:

$$[(1+2+\cdots+n)+(2+3+\cdots+n)+\cdots+n] + [1+(1+2)+\cdots+(1+2+\cdots+(n-1))] = n(1+2+\cdots+n).$$

21. 试用前面的三个结果证明:

$$1^2+2^2+\cdots+n^2=\left[n-\frac{1}{3}(n-1)\right][1+2+\cdots+n].$$

22. 证明两个相邻的三角形数的平方差是一个立方数.(将平方差分解成和与差的积,并利用练习 17 的结果.)(这个结果选自约丹努斯的《算术》. 同凯拉吉及莱维本热尔松一样,约丹努斯利用这个结果来证明,从 1 到 n 的整数的立方和等于第 n 个三角形数的平方.)

23. 回忆约丹努斯在《算术》的命题 IX – 70 中,如何用帕斯卡三角形确定连比例数列. 亦即,开始是数列 1,1,1,1…,先导出数列 1,2,4,8,…,再导出数列 1,3,9,27,…. 现用相同方法先导出数列 1,4,16,64,…,再把这个结果的一般情形公式化后用数学归纳法证之.

24. 用约丹努斯命题 IX – 70 中的相同方法,由数列 8,4,2,1 根据帕斯卡三角形构造连比例数列. 先导出数列 8,12,18,27,这个几何数列的公比是 3/2. 再导出其一般形式.

斐波那契的问题

25. 狮,豹,熊吃掉一只羊分别需要 4,5,6 小时,这三个动物一起吞没一只羊需要多久?(先假定答案是 60,即 4,5,6 的最小公倍数.)

26. 斐波那契数列可用递推法则给出,$F_0=F_1=1$ 且 $F_n=F_{n-1}+F_{n-2}$. 证明:

$$F_{n+1}\cdot F_{n-1}=F_n^2-(-1)^n, \text{且} \lim_{n\to\infty}\frac{F_n}{F_{n-1}}=\frac{1+\sqrt{5}}{2}.$$

27. 某数除以 2,3,4,5,6,7 分别剩余 1,2,3,4,5,0,求该数.(斐波那契可能从伊本海塞姆著作中发现这个问题的.)

28. 两数相差为 5,大数乘以$\sqrt{8}$ 等于小数乘以$\sqrt{10}$. 求这两数.

29. 证明莱昂那多的"和谐数"都能被 24 整除.

30. 选自《平方数书》:求一个平方数,无论加上或者减去它的平方根,其结果仍是平方数.(用现代的符号表示,即求 X,Y,Z,使 $X^2+X=Z^2$ 且 $X^2-X=Y^2$. 莱昂那多使用了和谐数 24 来求解 $a^2+24=b^2, a^2-24=c^2$,并用 24 导出了全部结果.)

约丹努斯《论给定的数》中的问题

31. 如果某一给定的数被分成两部分,这两部分的积与差的和是已知的,那么每一部分可被确定. 亦即,解方程组 $x+y=a, xy+x-y=b$. 用约丹努斯的例子,取 $a=9$ 且 $b=21$.

32. 如果某一给定的数被分成两部分,且分别除以另外两个给定的数后其和是已知的,那么每一部分可被确定. 亦即,解方程组 $x+y=a, x/b+y/c=d$. 约丹努斯取 $a=10, b=3, c=2$ 且 $d=4$.

33. 如果某一给定的数被分成两部分,且它们的平方积是已知的,那么每一部分可被确定. 约丹努斯的例子是 $x+y=9, x^2y^2=324$.

牛津大学和巴黎大学的问题

34. 根据奥雷姆的技巧,计算比率(3 : 2) 除以比率$(2:1)^{1/3}$.

35. 试说明,如果对从 2 : 1 到 101 : 1 的 100 个整数比率作比较,总是较大的比率比较小的,则有 4950 种情况,但确实只有 25 种情况具有有理指数.

36. 利用平均速度定理来说明,如果将时间区间四等分,则每一间隔内的位移之比为 1 : 3 : 5 : 7. 如果将时间区间 n

等分,则一般的结果是什么?并给出证明.

37. 选自奥雷姆《论质量与运动的形态幅度》:用几何方法证明下面级数的和等于 96

$$48\cdot 1+48\cdot\frac{1}{4}\cdot 2+48\cdot\left(\frac{1}{4}\right)^{2}\cdot 4+\cdots+48\cdot\left(\frac{1}{4}\right)^{n}2^{n}+\cdots.$$

38. 解出奥雷姆的下述问题:

将线段 AB(表示时间,长度为 1)按固定的比率 2:1 无限划分;即,第一部分是二分之一,第二部分是四分之一,第三部分是八分之一,等等.在第一部分有一个匀速运动(速率恒为 1),在第二部分有一个匀加速运动(速率从 1 到 2),在第三部分是匀速为 2,在第四部分是从 2 到 4 的匀加速,等等(图 8.17).试说明总位移是 7/4.

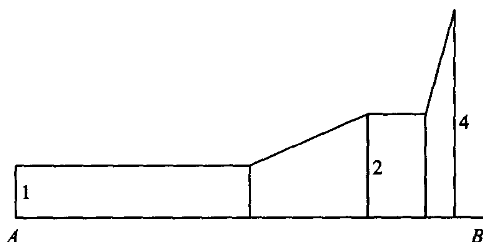


图 8.17 奥雷姆的一个问题.

39. 试验证奥雷姆的结论:调和级数(1 + 1/2 + 1/3 + 1/4 + …) 的和为无穷大.

讨论题

40. 指出解决复活节问题所需要的数学知识.教会争论的结果是什么?复活节的日期在当今是如何确定的?(注意,罗马天主教的方法与东正教不同.)
41. 试比较莱维本热尔松与凯拉吉关于"数学归纳法" 的使用,并讨论.
42. 试用莱维本热尔松的例子来演示归纳证明,书面完成.
43. 试用伊本艾兹拉及莱维本热尔松的方法,针对某些基本的组合学公式作出推导,书面完成.
44. 试用几何方法证明某些无穷级数的和,可以把奥雷姆的例子一般化,书面完成.

文献和注解

最好的关于中世纪欧洲数学的综合性资料有:Marshall Clagett, *Mathematics and Its Applications to Science and Natural Philosophy in the Middle Ages*(Cambridge: Cambridge University Press, 1987);David C. Lindberg, ed., *Science in the Middle Ages*(Chicago: University of Chicago Press, 1978).特别是后一著作的第 5 章论数学(Michael S. Mahoney 撰)和第 7 章论运动科学(John E. Murdoch 和 Edith D. Sylla 撰)提供了很好的综述.最佳的将原始资料译成英文的选集是:Edward Grant, ed., *A Source Book in Medieval Science*(Cambridge: Harvard University Press, 1974);另一本主要涉及力学方面的原始资料选集是 Marshall Clagett, *The Science of Mechanics in the Middle Ages*(Madison: University of Wisconsin Press, 1961).关于中世纪数学家的基本传记性资料可查阅:GeorgeSarton, *Introduction to the History of Science*(Huntington, New York: Robert E. Krieger, 1975),尤其是 vol. II and III.最后,一本不仅含有欧洲,还有伊斯兰,中国,印度的中世纪数学的系统教程是 Adolf P. Juschkewitsch, *Geschichte der Mathematik in Mittelalter*(Leipzig: Teubner, 1964),系 1961 年的俄文著作的德文译本.

1. Maximilian Curtze, "Der Liber Embadorum des Abraham bar Chijja Savasorda in der Ubersetzung des Plato von Tivoli," *Abhandlungen zur Geschichte der mathematischen Wissenschaften* 12(1902), 1 – 183, p. 11.该文提供了拉丁版的德文译本.目前还没有直接译自希伯来文原著的现代译本.
2. Leonardo Pisano Fibonacci, *The Book of Squares*, edited and translated by L. E. Sigler(Boston: Academic Press, 1987), p. 3.

Sigler 评注和翻译了莱昂那多的原著,并将其中许多内容译为现代符号.

3. St. Augustine, *City of God*, XI, 30, 引自 Wisdom of Solomon, 11:20.

4. Curtze, "Liber Embadorum," p. 35.

5. 对于圣维克多的于格的原著的翻译和评注,参见 Frederick A. Homann, S. J., *Practical Geometry, Attributed to Hugh of St. Victor* (Milwaukee: Marquette University Press, 1991).

6. Stephen Victor, *Practical Geometry in the High Middle Ages. Artis Cuiuslibet Consummatio and the Pratike de Geometrie* (Philadelphia: American Philosophical Society, 1979), pp. 109 – 111. 这本书包含许多中世纪实用几何传统的综合资料,并翻译了书名中的两部原著. 另一部实用几何著作在 H. L. Busard, "The *Practica Geometriae* of Dominicus de Calvasio," (*Archive for History of Exact Sciences* 2(1965), 520 – 575) 一文中作了详细讨论. 对其它类似著作的讨论见 Gillian Evans, "The 'Sub-Euclidean' Geometry of the Earlier Middle Ages, up to the Mid-Twelfth Century," *Archive for History of Exact Sciences* 16(1976), 105 – 118.

7. 同上, p. 221.

8. 同上, p. 295.

9. Isidor Kalisch 编译 *The Sepher Yezirah* (Gillette, N. J.: Heptangle Books, 1987), p. 23.

10. 引自 Nachum L. Rabinovitch, *Probability and Statistical Inference in Ancient and Medieval Jewish Literature* (Toronto: University of Toronto Press, 1973), p. 144. 这本书对古代和中世纪各地的犹太教学者关于概率和统计的各种基本思想的萌芽提供了有趣的资料. 一本更早的关于犹太数学的著作是 M. Steinshneider, *Mathematik bei den Juden* (Hildesheim: Georg Olms, 1964), 系 1893—1901 年间原始论文的重印本.

11. J. Ginsburg, "Rabbi ben Ezra on Permutations and Combinations," *The Mathematics Teacher* 15(1922), 347 – 356, p. 351.

12. Gerson Lange, *Sefer Maasei Choscheb. Die Praxis der Rechners. Ein hebräisch-arithmetisches Werk des Levi ben Gerschom aus dem Jahre* 1321 (Frandfurt: Louis Golde, 1909), p. 1. 这本书含有莱维本热尔松关于组合学的希伯来文原著及德文译本. 随后的命题均译自德文. 关于莱维的归纳法思想,更详细的讨论参见 Nachum L. Rabinovitch, "Rabbi Levi ben Gershon and the Origins of Mathematical Induction," *Archive for History of Exact Sciences* 6(1970), 237 – 248.

13. 同上, p. 8.

14. 同上, p. 49.

15. 同上, p. 51.

16. B. Boncompagni, ed., *Scritti di Leonardo Pisano* (Rome: Tipografia delle scienze matematiche e fisiche, 1857—1862), vol. 1, p. 2. 这本书是莱昂那多全集的标准版. 第一卷含有《算盘书》. 我们所讨论的莱昂那多的各种问题均取自这个版本.

17. 同上, p. 283.

18. Leonardo Pisano, *The Book of Squares*, p. 3.

19. 关于哈津的资料,参见 Roshdi Rashed, *The Development of Arabic Mathematics*, Chapter IV. 关于两人著作的关系的更详细介绍,参见 Roshdi Rashed, "Fibonacci etles mathématiques arabes," 载 *Le scienze alla corte di Federico II* (Paris: Brepols, 1994), pp. 145 – 160.

20. 这封信引自 Jens Høyrup, "Jordanus de Nemore, 13th Century Mathematical Innovator: an Essay on Intellectual Context, Achievement, and Failure," *Archive for History of Exact Sciences* 38(1988), 307 – 363. 这篇论文对约丹努斯的著作做了广泛的考察. Wilbur Knorr 的文章 "On a Medieval Circle Quadrature: *De circulo quadrando*," *Historia Mathematica* 18(1991), 107 – 128 作出了新的尝试,将约丹努斯放到 13 世纪早期整个巴黎的背景中来论述.

21. Barnabas Hughes, "The Arithmetical Triangle of Jordanus de Nemore," *Historia Mathematica* 16(1989), 213 – 223. H. L. L. Bursard, *Jordanus de Nemore, De Elementis Arithmetice Artis* (Stuttgart: Franz Steiner Verlag, 1991) 则提供了整个 *Arithmetica* 的校勘版,并给出了英文意译.

22. Barnabas Hughes, *Jordanus de Nemore: De numeris datis* (Berkeley: University of California Press, 1981), p. 57. 这本书将约丹努斯的代数著作从拉丁版译为英文,各命题附有现代符号表示,还讨论了原著的资料来源. 但本书下面几页引征的命题则是直接译自拉丁文,以使读者能比从 Hughes 的现代翻译更多地领略约丹努斯原作的风味.

23. 这些定义引自 John North, *Richard of Wallingford. An Edition of His Writings With Introductions, English Translations and Commentary*(Oxford: Clarendon Press, 1976), p.59. 这套三卷本的校勘本提供了理查德主要数学与天文著作的英译与评述.

24. Edward Grant, *Nicole Oresme: De proportionibus proportionum and Ad pauca respicientes*(Madison: University of Wisconsin press, 1966), p.161. 这本书不仅含有奥雷姆《论比率之比率》的拉丁版及完整的英文译本,还对奥雷姆的思想背景作了广泛分析.

25. 同上, pp.161 – 163.

26. Clagett, *The Science of Mechanics*, p.230. 关于布雷德沃丁及其同时代人的工作的更详尽的介绍可参阅 A.G. Molland, "The Geometrical Background to the 'Merton School': An Exploration into the Application of Mathematics to Natural Philosophy in the Fourteenth Century," *British Journal for the History of Science* 4(1968), 108 – 125.

27. 同上, pp.230 – 231.

28. Grant, *A Source Book in Medieval Science*, p.238. 亦见, Clagett, *The Science of Mechanic*, p.236. 对海特斯伯里工作的更详尽的研究参见 Curtis Wilson, *William Heytesbury: Medieval Logic and the Rise of Mathematical Physics*(Madison: University of Wisconsin Press, 1956).

29. Grant, *Source Book*, p.238; Clagett, *The Science of Mechanics*, p.237.

30. Grant, *Source Book*, p.239; Clagett, *The Science of Mechanics*, p.271.

31. Marshall Clagett, *Nicole Oresme and the Medieval Geometry of Qualities and Motions. A Treatise on the Uniformity and Difformity of Intensities Known as Tractatus de configurationibus qualitatum*(Madison: University of Wisconsin Press, 1968), pp.165 – 167. 除了拉丁原著及完整的英译本外, 这本书还有一个详细的评论及某些相关著作的译文.

32. 同上, p.167. 关于用线表示函数关系的研究, 参见 Edith Sylla, "Medieval Concepts of the Latitude of Forms: The Oxford Calculators," *Archives d'Histoire Doctrinal et Littérarie du Moyen Âge* 40(1973), 223 – 283.

33. 同上, p.193.

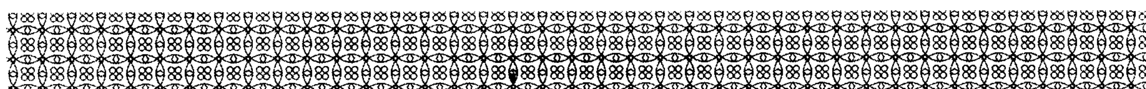
34. 同上, p.193.

35. 同上, p.415.

36. David Singmaster, "Some early sources in recreational mathematics," 见 Cynthia Hay 主编, *Mathematics from Manuscript to Print*: 1300 – 1600(Oxford: Clarendon Press, 1988), p.199.

中世纪数学概览

480—524	博依西斯(Boethius)	希腊著作的拉丁文译本
560—636	塞维利亚的伊西多雷(Isidore of Seville)	四艺的拉丁文教科书
735—804	约克的阿尔昆(Alcuin of York)	算术问题
945—1003	奥里亚克的热尔贝(Gerbert d'Aurillac)	算盘
1075—1164	巴思的阿德拉德(Adelard of Bath)	翻译
12 世纪初	塞维利亚的约翰(John of Seville)	翻译
12 世纪初	阿勃拉汉·巴·希亚(Abraham bar Hiyya)	几何和三角学
12 世纪初	蒂沃利的普拉托(Plato of Tivoli)	翻译
1090—1167	亚伯拉罕·依本艾兹拉(Abraham ibn Ezra)	组合论
12 世纪中	切斯特的罗伯特(Robert of Chester)	翻译
1114—1187	克雷莫纳的杰拉德(Gerard of Cremona)	翻译
1170—1240	比萨的莱昂那多(即斐波那契)(Leonardo of Pisa)	算术、代数、几何
13 世纪初	约丹努斯(Jordanus de Nemore)	算术、代数
1225—1286	莫贝克的威廉(Wilhelm of Moerbeke)	翻译
1288—1344	莱维·本热尔松(Levi ben Gerson)	三角学、组合论、归纳法
1291—1336	沃林福德的理查德(Richard of Wallingford)	三角学、比例
1295—1349	托马斯·布雷德沃丁(Thomas Bradwardine)	运动学
14 世纪初	威廉·海特斯伯利(William Heytesbury)	运动学
1320—1387	尼古拉·奥雷姆(Nicole Oresme)	运动学、指数、图像



插入章　世界各地的数学

较详细地研究了大约到1300年的中国、印度、伊斯兰世界，及欧洲的数学之后，有必要对这些地区在当时的数学做出比较．由于这些相互独立的文化中有着许多共同的思想，我们将对其中一些思想从一种文化传向另一种文化的可能性作讨论，并提出这样的疑问：为什么现代数学起源于欧洲，而非别的地区？另外，在世界的其它地区，过去的数学水平如何？就我们目前所知，第二个问题难以详述，但我们将对美洲、非洲，以及太平洋地区的数学知识作一个简介．

Ⅰ.1　14世纪转折时期的数学

首先讨论几何学．实用几何学的研究——亦即，田地的测量，距离和高度的确定，体积的计算，等等——在上述四个地区，基本上采用着相同的技术．那里的人们都知道怎样计算面积和体积，在讨论直角三角形时，都知道并使用了毕达哥拉斯定理．甚至测定塔高的技术也类似．

对于理论几何学来说，正是在伊斯兰世界中，传统的希腊几何学遗产才得以保存和研究，并在某些方面有所发展．正是在那里，有些问题被提出并被解答，诸如某些立体的精确体积以及重心的位置，其中既使用了直接推断法来得到答案，又使用了穷举技巧来给出证明．正是在那里，有些问题被提出并试图解答欧几里得的平行线公设．正是在那里，有些问题被提出并发展了希腊人区别数和量的传统思想．正是在那里，希腊人从所规定的公理出发来证题的思想，得到了最充分的理解及发展．

虽然在欧洲至少可见到欧几里得《原本》原著的一种版本，但只是在12和13世纪大量出现的译本的刺激下，在14世纪初，对欧几里得和其他希腊几何学家的兴趣才开始得到复兴．不过证明的思想虽然保留了下来，却没有出现关于理论几何的新著．而印度和中国，就目前所知，并没有接触到经典的希腊几何学，但这并不意味着他们没有证明的概念．在中国数学家及其注释者的著作中，一直存在着对结果的推导．然而，这些推导并非基于严格规定的公理．从另一个角度说，它们是逻辑推理的举例．在印度，推导和解释甚至更缺乏，但仍能看出数学家们某种程度上在思考其理由，为什么一定的过程便能给出正确的答案．

相关于几何学的是三角学，作为天文学研究的一部分，它起源于古希腊世界．到1300年时，三角学已在印度、伊斯兰、和欧洲处于实际应用中，一般都是为了研究天体的目的．三角学在传播中曾

被修改和扩充,这三个文明中对天体有专爱者一般都精通三角学.而中国似乎缺乏三角学知识,虽然印度学者在8世纪时曾将其基本内容传入过中国.也许三角学对中国人来说,在他们的天文历法计算中没有用处吧.

另一方面,在代数学的某些领域,中国人最早发现了一些后来才在别的地区被应用的技能.例如,他们早已发现了解线性方程组的有效方法.到14世纪时,他们已将其早期的求根技术,其中涉及到帕斯卡三角的使用,发展成解任意次多项式方程的详细过程.他们还奠定了我们今天所说的中国剩余定理的基础,即解联立一次同余式组的方法步骤.

印度人也解决了一次同余式问题,其方法与中国人的不同,但仍然应用了欧几里得算法.然而,印度学者更自豪的是他们发展了解二次不定方程的技术,亦即今天所说的佩尔(Pell)方程.虽然有证据表明希腊人曾研究过这些方程的某些简单个例,印度人所发现的一般方法直到18世纪,才在欧洲得到研究.印度的数学家对解二次方程的标准方法也很熟悉,由于没有证据表明他们是如何发现这个方法的,我们就不能肯定到底是他们独立地发现了这个技能,还是从古代巴比伦人那里吸收了它.第三种可能是他们从丢番图的著作中学会了它,而丢番图至少不是直接地从巴比伦人那里知道了这些方法.

对伊斯兰,当然有丰富证据说明其对代数的兴趣.伊斯兰数学家不仅曾非常详细地研究了二次方程,对解题时的各种代数过程给出了几何验证,而且还研究了三次方程.对这些方程,伊斯兰数学家发现了一种涉及到圆锥曲线的解法,并对这些方程的根与系数的关系有了一定程度的了解.另外,他们也知道解多项式方程的数值方法,其类似于中国人的方法,并完全基于帕斯卡三角形.

在相关二项式定理及组合学研究中,帕斯卡三角形也出现在伊斯兰数学中.那些论述这两个方面的帕斯卡三角形的伊斯兰数学家采用归纳法,还发现了完全类似于现代证法的证明技术.欧洲的莱维(Levi ben Gerson)进一步发展了这些技术.还有,伊斯兰代数学家推进了代数表示的方法,特别是涉及到根式,并因此而开始了否定希腊人将数与量进行区分的过程.

直到14世纪的交替时期,代数技能才开始出现在欧洲.那些技能明显基于伊斯兰著作,虽然约丹努斯考虑问题的观点有点不同.他还引入了一套符号体系,而伊斯兰代数中却完全没有,且在形式上不同于印度和中国.另一方面,这个时期的欧洲代数,和它的伊斯兰相应方面一样,根本没有考虑负数.然而,印度和中国在计算时,却能熟练应用负数,即使他们在数学问题的解答中对其使用仍很犹豫.

这个时期出现在欧洲,却明显没有出现在别的地区的一个数学论题是,相关于运动的复杂思想.显然只有在欧洲,数学家在思考瞬时速度的数学意义,并发现了平均速度法则.这样,就播下了在大约三个世纪后,成为微积分的一个分支的种子.

似乎在14世纪的交替时期,世界上的这四个地区的数学水平是能够比较的.虽然有些技能只在一种文化中存在,仍有很多数学思想和方法为两个或更多地区共有.那么问题就出来了,那些思想到底是在这四个地区独立发展的,还是在他们之间传播的?

对于某些思想,传播的线路是清楚的.三角学是从希腊传到印度,再到伊斯兰,后又回到欧洲,并伴有各文化的烙印,以适应各自的要求.还有十进制的位值系统,发端于中国或印度(或其边界地区),在8世纪时传到巴格达,后在11和12世纪时传到欧洲(通过意大利和西班牙).

但对其它一些共同思想,情况就不甚清楚.例如,三角学中的正切函数,在用阴影的长度计算太阳的高度时很有用.最早的正切函数表在8世纪早期出现在中国.一行对这种思想的发展也许是由于印度人对正弦的计算的帮助.下一个所发现的正切表是在伊斯兰.这个思想是由公元751年在塔拉斯河战争中被俘的中国技术人员带到那里的吗?这次战争建立了伊斯兰在中亚西部的霸权.虽然

正切表以及花拉子米的天文表的译本在12世纪早期被传入欧洲,但正切函数并未出现在早期欧洲人自己的三角学著作中.

帕斯卡三角形的情况如何?它在11世纪早期出现在伊斯兰,并在该世纪中期在中国发现.是传播的吗?这个时期伊斯兰和中国沿着著名的丝绸之路确有交流.而且这个时期比鲁尼就在伽兹纳的默汗穆德苏丹朝廷中研究印度文化,而印度和中国一直有着某种联系,特别是通过佛教.后来到13世纪早期时,帕斯卡三角形怎样传到了欧洲?欧洲数学家是通过今天我们所不知道的阿拉伯手稿学到这种数字的排列的吗?或者这个知识是由在伊斯兰和欧洲之间旅行的学者传播的吗?

对于四种文化中都存在的确定高度和距离的方法来说,据记载,有两处观测点的基本方法最早在公元3世纪时已出现在中国.到13世纪时,这种方法才在欧洲被应用.类似的趣题还有百禽问题,抽空(或充满)容器问题,它们早已出现在中国和伊斯兰,后又出现在中世纪早期欧洲人的著作中.这些问题是传播的吗?如果是,又是怎样传播的?使人再次想起丝绸之路可作为传播途径.更具体地联想到,公元9世纪时,曾有一批犹太商人(Radhanites)定期地从法国南部,经大马士革和印度,游历中国.他们曾经学到过中国数学并带回到欧洲,或途经的某地吗?或相反地,他们曾将伊斯兰或印度的数学带到过中国吗?这个问题甚至和雅各标杆(the Jacob's Staff)问题有关系,这种观测仪器在欧洲最早是在14世纪早期时,由莱维所描述,而中国11世纪早已有之.那么,它是犹太商人从中国带入的吗?

关于传播与交流的许多问题的答案目前只能去推测.其证据有待于发现.不管怎样,这些共同的数学思想都是为了满足各种文明的数学需要.

另一个广泛被讨论的问题是,为什么现代数学(以及现代科学)出现在西欧,而不是伊斯兰世界或印度或中国?由于四种文明的技术成就在1300年左右时是相近的,许多学者便从宗教信仰和文化背景的角度寻求其答案.[1] 例如,在中国,实质上仅有一所"大学".而且它是帝制管理中官僚政治的产物.因此,除非帝制管理部门鼓励数学发展——这几乎未曾发生——就没有地方使有一定数学修养的人能够发展其思想.正如我们所知,中国数学家从时间和地理位置上讲,都是孤立的,也许他们中的一些人并不知道别人的存在.因此,虽然有过一些新思想的发现,但政府却从未有过足够的热情来激励新思想.事实上,中国的教育通常致力于对古代经典的背诵与评注.

另一方面,伊斯兰有着大量的高级研究组织以及大量对数学有兴趣的学者.因此,数学传统得以发展,一个人甚至可以建立数学"学校",形成一组学者来用相关的技能研究类似的问题.那么,不可思议的是,为什么伊斯兰没有找到通往现代数学的突破口?譬如说,为什么伊斯兰数学家没有发现微积分或日心天文学?事实上,正如我们前面所说,伊斯兰数学在13世纪之后,遭受了一段时期的衰退,早期的重要思想被丢失.

虽然我们并不肯定,但似乎关键因素之一是,即使数学在伊斯兰高度发展时期,数学中那些比基本的算术更高级的部分仍被归为"外在的科学",这是相对于"宗教科学"而言,宗教科学包含宗教法律和神学推理.伊斯兰的宗教领袖认为,外在的科学对宗教信仰具有潜在的颠覆作用,对今世或来世的生活需要来说,是多余的.虽然最早的伊斯兰领袖鼓励对外在科学的研究,但在数世纪之后,当更为保守的宗教领袖登上前台时,对这种研究的支持在减少.比鲁尼甚至在11世纪时已经认识到了这一点:

> 学海无涯,如果人们能尊重知识和人才,使科学处于支配地位并为大家所青睐,科学发展的速度便会加快.而这首先是统治者、即王公们的责任和义务.只有他们才能使学者免去日常生活所需的烦恼,激发学者去获得更大声望,而对声望的追求是人性的精髓.然而,现今的情况完全相反,因此,不可能有新科学或新的创造性研究的出现.我们所掌握的

知识只不过是过去辉煌时代的东西而已.[2]

虽然在比鲁尼之后，伊斯兰在数学上仍有某些成就，但其发展的速度在减缓.尽管在整个伊斯兰世界仍保留着高级研究机构(madrasas)，但它们越来越专注于伊斯兰法律的教学.如果学者要想讲授那些外在的学问，必须经过因循守旧者的法定裁决，说明所讲内容不会有害于伊斯兰的宗教信条.

有趣的是，欧洲的天主教领袖也曾颁布教令，禁止某些课题的讲授.事实上，曾几次正式禁止讲授亚里士多德著作中明显与教义相抵触的内容.然而，巴黎以及其它地方的大学中的学者们，似乎并未理会教会的教令.欧洲的大学与伊斯兰的高级研究机构(madrasas)不同，它们是法人团体，在法律上具有明确的自治权.如果教学人员决定要讨论科学问题，并发展新数学思想的话，教会的领袖通常是难以禁止的.因此，发展现代数学，当然还有现代科学的通道，在欧洲就打开了.

这并不否定伊斯兰成就的作用.相反地，我们对12和13世纪期间传入欧洲的伊斯兰数学成就的各个方面，均已有说明.在数学的某些方面，欧洲的早期思想直接来源于伊斯兰的成就，或伊斯兰对希腊及印度成就的修改.在随后章节中，我们会详述伊斯兰对欧洲数学的某些影响.然而，14世纪之后，数学史的焦点主要在欧洲.因此，在本书后面的若干章节中，将着眼于在那里所产生的数学成就.

Ⅰ.2　美洲、非洲以及太平洋地区的数学

在那些与前面已经讨论过的四种主要的中世纪文明不同的社会中，也存在着数学思想.不幸地，大多数其它文明没有文字，故不可能得到书面记载.因此，关于这些社会的数学描述必须依据人类学家的理论及留传下来的实物.近年来已做出很多研究，但仍有许多问题未解决.此处只能就目前所知的几个社会中的数学思想作些简介.不过，附有一些参考书目，以便有兴趣者进一步探索.

我们首先讨论美洲的玛雅文明，玛雅人过去已有文字语言.玛雅文明繁荣于墨西哥南部，危地



图Ⅰ.1　刻有两位数学家的玛雅陶器图(公元750年).左侧是男性，右上角是女性.其身份由腋下数符饰带鉴定.

马拉、伯利兹,及洪都拉斯,在公元3和9世纪期间达到其辉煌时期.后来,玛雅人受到墨西哥其他民族的影响,而且他们的许多文化中心已被毁灭.然而,即使在16世纪早期,当西班牙人到达时,浓厚的玛雅文化仍保存着.虽然西班牙人统治着玛雅人,他们却未能完全毁灭玛雅文化.今天,仍有接近250万人能讲玛雅语,在某些方面保持着他们的古代生活方式.

像许多古代文明一样,玛雅人有一个教士阶层,其研究数学和天文学,并管理历法(图I.1).教士的记录会写在一种树皮纸上,或刻在石碑上.不幸地,西班牙统治者毁灭了绝大多数这样的证据,故流传至今者极少见.由于现今的玛雅人不懂古代的象形文字,故破解那些保存下来的少量资料,是非常复杂和漫长的工作,特别是那个叫做"德累斯顿抄本(Dresden codex)"的文物(以拥有它的图书馆而命名),它的日期被定为12世纪,涉及到玛雅的历法(图I.2).然而,学者们现已弄清了玛雅历法和记数系统的基础.不过,那些记录仅提供了计算结果,而没有计算方法.因此,对玛雅数学的描述,有时只能是推测性的.

像巴比伦人一样,玛雅的记数系统是一种混合系统.基本上采用的是20进位的位值系统,但表示数字的符号只有两个,其中一个点(·)表示1,一条线(——)表示5.要表示19以内的数字,就必须对其进行适当的组合,例如,用 $\overset{\cdots}{=}$ 表示8,用 $\overset{\cdot\cdot}{\equiv}$ 表示17.对于大于19的数字,就要用到位值记数系统.第一位的位值是单位1,第二位是20,第三位是400,等等.但是,和巴比伦人不同的是,玛雅人有一个表示0的符号,即⊖,用它来指出一个"空"位.玛雅数字一般是竖写式,最上面的位值最大,但为了方便,我们将采用横写式,这与表示巴比伦数字时的约定一样.因此,3,5就表示 $3\times20+5$,即65.



图I.2　德累斯顿古抄本,德意志民主共和国邮票.

为了历法上的应用,玛雅人稍微修改了他们的位值记数系统,用从下算起的第三位表示360,而不是400,至于其它各位,仍然表示其前面位值的20倍.我们后面将采用这种系统,因为在历法计算中,玛雅人通常使用这种数字.例如,在这种历法读数系统中,2,3,5表示 $2\times360+3\times20+5$,即785,而2,0,12,15表示 $2\times7200+12\times20+15$,即14655.

对于巴比伦人来说,有大量证据说明他们所采用的计算方法.那么,玛雅人在他们的位值系统中,是怎样计算的?很不幸,所保存下来的玛雅记载都是计算结果,虽然几乎全部都是正确的,但却没有任何关于方法本身的记录.可以推测,对于加减法,玛雅人曾用过某种形式的计算板来堆积各个数位上的点和线,并在超过20时便移到下一位.对于乘法来说,大概所需要知道的是三个事实:$1\times1=1$,$1\times5=5$ 和 $5\times5=1,5$.要做具体的乘法运算的话,只需要再知道分配律和一种掌握位置关系的方法,当然,如果知道 19×19 以内的乘法表,会使运算方便些.然而,没有发现关于这些乘法表的证据.

玛雅教士对数字的计算主要用于计算日历.要理解这些,就得首先了解玛雅的日历制度.玛雅人过去同时使用着两种不同的历法.第一种历法,一个历年一共有260天,由1–13这十三个数字和另外二十个专用名按顺序循环相配而成.亦即,一年中的任一天都可用一个数组 (t,v) 表之,其中 t 是1–13中的一个日数,v 是二十个日名之一.例如,因为二十个日名的前两个分别是伊米克司(Imix)和伊克(Ik),所以,日历1伊米克司可表示为(1,1),而日历5伊克可表为(5,2).第二种历法是一年365天.这种历年被分成18个月加另外的5天,其中每月20天.对于我们下面的讨论来说,只用一个数字 y,就足以反映这种历年中的某一天.例如,因为磨安(Muan)是这种日历的第十五个月,所以,磨安月的第二天就可表示成 $y=282$.这三个周期,亦即13个日数,20个日名,以及第二种日历年的365天是相互独立地变化的,因此,三元数组 (t,v,y) 的总周期为 $13\cdot20\cdot73=18980$ 天,或

52 个第二种历年，或 73 个第一种历年. 这个总周期通常称之为“日历圈”(Galendar-round).

玛雅人需要解决的两个基本的日历问题是，首先，(以三元数组的形式) 给定一个日期和一段时间之后，如何确定新的日期?另外，(以三元数组的形式) 给定两个日期之后，如何确定它们之间的最短时间间隔? 如果某一段日期用 20 进制的日历符号表示为 m,n,p,q,r，其中 $0 \leqslant m,n,p,r \leqslant 19$，$0 \leqslant q \leqslant 17$，第一个问题用现代符号可表述为，给定一个初始日期$(t_0,v_0,y_0)$，确定 m,n,p,q,r 天之后的新日期(t,v,y) 是什么. 可以给出如下的解答：

$$t = t_0 - m - 2n - 4p + 7q + r \pmod{13},$$
$$v = v_0 + r \pmod{20},$$
$$y = y_0 + 190m - 100n - 5p + 20q + r \pmod{365}.$$

[第一个方程的系数是因为 $20 \equiv 7 \pmod{13}$，$18 \times 20 \equiv -4 \pmod{13}$，$20 \times 18 \times 20 \equiv -2 \pmod{13}$，和 $20 \times 20 \times 18 \times 20 \equiv -1 \pmod{13}$.] 例如，如果所给的日期是(4,15,120)，在 0,2,5,11,18 天之后，新的日期为(10,13,133). 对于第二个日历问题，首先需要在三个分量各自的周期内，分别确定两个日期之间的最短间隔，第二步，把前两个计算结果合并起来，确定在日历周内的最短间隔，第三步，再把这个值与第三个计算结果合并起来，确定在日历圈内的天数.

没有证据能表明，玛雅教士在解决这些问题时所用的确切方法. 但是，几乎可以肯定，他们所采用的算法应该类似于上述代数公式，当然，代数符号本身是不会有的. 总之，教士们一定能够解决这些历日计算中的问题，以便给玛雅官方提供出精确的日子，来庆祝节日、典祀，种植玉米，或完成任何管理玛雅王国所需的其它任务(图 I.3).[4]



图 I.3 玛雅天文台，位于墨西哥尤卡坦半岛的奇陈 - 伊查(Chichen Itza) 市.

在距玛雅文化中心大约 2000 英里的南方，有另一个主要的文明，印卡人(Inca)，大约有 400 万人口. 印卡人没有书面语言，但却拥有一种“绳结语”(quipus). 绳结语是印卡官方控制其领土的工具. 他们日常必须接送大量信息，包括库存、税收、某项工程所需人数等细目. 这些信息被译成绳结语，并由信差送往目的地(图 I.4). 当然，绳结语应该简练，故其首都库斯科(Cuzco) 有专人培训如何设计制作绳结语.[5]

绳结语由各种颜色与形式的绳结组成，其中颜色、绳子的布置，各条绳上的绳结，绳结的布置以及绳结间的距离，都被用于表示所记录的数据的含意. 绳结语一般都有一条较粗的主绳，其上系着若干细绳(叫做悬挂线)，而悬挂线可能还系着别的细绳(叫做补充线). 用一系列绳结把数据记录在这些细绳上(并非主绳). 绳结被分组束在一起，各组之间隔有一定距离. 表示数字时，采用 10 进位的位值系统，离主绳近者，其位值最大. 目前在绳结语中，所发现的最大的数字是 97357. 零一般用特别宽的间隙来表示(图 Ⅰ.5). 一般来说，绳结语中没有计算过程，仅记录着结果. 这些结果所依赖的运算过程，可能是在某种算板上进行的. 另外，绳结语的结构可以看作现代术语中的某种特殊的“树图”(见第 18 章). 其设计者肯定会思考一些“树图” 中的基本问题，以便绳结语更有效和易识别.[6]



图 I.4 传送绳结语的印卡信差.

在印卡文明中，像玛雅一样，有一个专业的“数学家” 阶层，来探讨其文化中的数学问题. 但是，在下面所要讨论的其它文化中，这样的阶层是不存在的. 那些人的头脑中并无“数学” 这样的分类，然而，其文化的某些方面具有明显的数学特征. 数学思想也是他们日常生活、耕作、建筑、礼拜所需

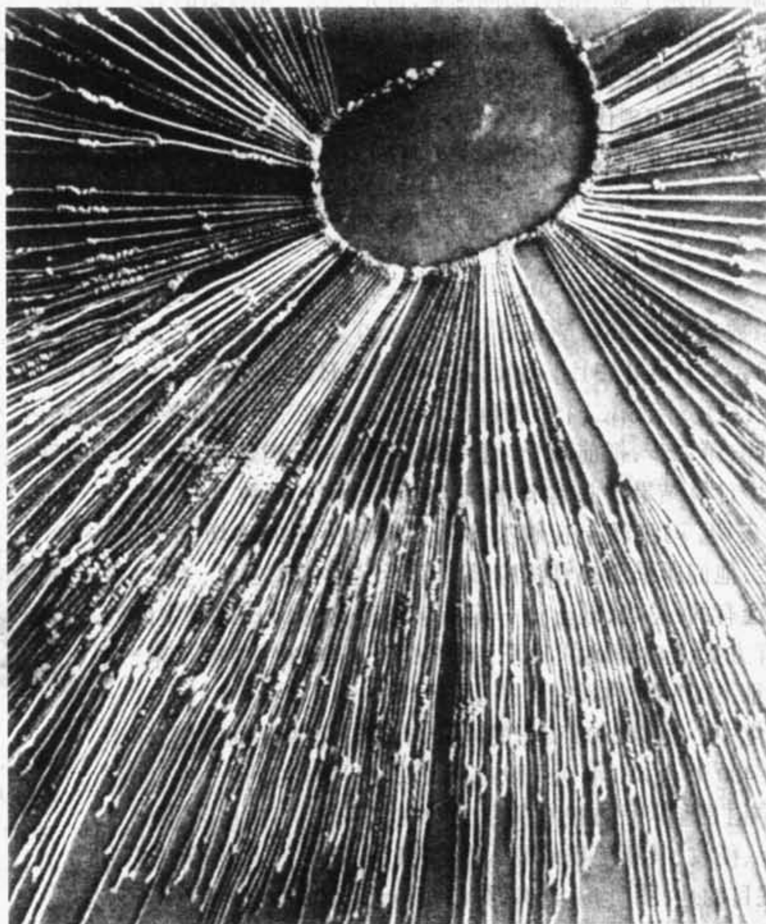


图 I.5　绳结语，秘鲁首都利马博物馆.

的一个基本方面.现在，把少数民族部落中的这种数学，通常称之为“民族数学”，对这种数学的研究可以使我们了解到，对于这样的部族来说，数学思想的重要性.[7]

在现今美国的前哥伦布时代，最具代表性的是阿纳萨支(Anasazi)文明，生活在西南部的四角形(Four Corners)地区.在公元1000年后，他们的文明达到最辉煌时期，在此期间在各处建造了复杂的村落和举行仪式的建筑，其中最突出的位于科罗拉多州西南的弗德台地及新墨西哥州西北的查科谷(Chaco Canyon).在巨石阵(Stonehenge)神殿的观望线上，多处建筑中建有观望台，以确定某些重要的天文现象，包括冬至、夏至，等等.

虽然，阿纳萨支人没有留下有关他们的数学的物证，但我们可通过考古遗迹来推测其生活所需的数学.例如，四个基本方向是阿纳萨支宗教中的一个重要概念，其来源于有关该民族起源的神话传说.对阿纳萨支人来说，按照四个基本方向使他们的主要建筑排成一线是非常重要的，甚至也要那样来修路.从查科谷出发，向正北方向就修有一条很长的那样的大路，而不顾地形的障碍.在查科谷有一个直径60英尺，礼仪用的环形建筑，其顶部由四根支柱支撑，四根支柱完全按四个基本方向形成一个正方形的四个角.那么，问题是阿纳萨支人如何确定正北方向.由于他们知道太阳在每天

和每年的运动,故一种可能是,他们也采用了与一千年前古罗马观测者一样的技术,即,以一根柱子为中心画一个圆,再把一天中柱子阴影的端点曲线记录下来,并确定曲线与圆的两个交点.这两点的连线就是东西方向,而划出其垂直平分线就可确定南北方向(见图 5.1)[8].

还有其它的北美印第安人也建造了排成一行的建筑,甚至整个市区,从而显示出他们的天文和几何知识.例如,在伊利诺州的东圣路易有一个叫做加荷基的护堤,由密西西比人建于 900 到 1200 年期间,不仅显示出与重要天象的一致,还有详细的城市规划的迹象.类似地,由大平原印第安人建造的大盘驱魔轮(Bighorn Medicine Wheel) 和麋鹿山驱魔轮(Moose Mountain Medicine Wheel),可能是为了确定夏至,其中,第一个位于怀俄明州北部的梅迪辛山的顶部,可追溯到 2000 年前,第二个位于[加拿大] 萨克其万省的东南方,要近代得很多.

像北美印第安人一样,大多数过去的非洲文化没有文字记载,因此,不可能确切说明在其文化中的数学思想是何时及怎样被创造的.对史学家来说,更糟的是在撒哈拉以南的非洲几乎没有能反映其数学思想的文物.不过,在非洲南部的津巴布韦的尼安达(Nyanda) 南边 17 英里处,有一个石混结构的古建筑"大津巴布韦",现正在被详细研究,它可能建于 12 世纪.很显然,建造如此巨大建筑的帝国需要数学来满足管理及工程的需求,需要数学处理贸易、税收、日历方面的问题以使帝国能有效运转.类似地,中世纪西非国家的官僚们,包括加纳、马里,及桑海(Songhai),也和世界其它地方的同僚一样,需要数学.因为伊斯兰的影响曾渗透到西非,在[马里]廷巴克图从 14 世纪到 1600 年期间,曾有过一个伊斯兰大学,这个地区的学者可能面临过一些伊斯兰的数学.然而,我们没有关于这个时期和地区的数学的任何资料.

在更多的考古文物被发现之前,我们可以对 19 和 20 世纪中那些专门研究非洲人的人类学家的报告进行分析,并从中提炼和找出非洲人的数学思想.另外,我们也可对那些未曾受到殖民主义者影响的当地的实践活动进行研究.近年来在莫桑比克及非洲南部的周边地区就有一些这样的研究出现.[9] 不幸地,这些方法都不能使我们确定那些思想的出处及日期.

在扎伊尔的布松(Bushoong) 文化和安哥拉东北部的首魁(Tshokwe) 文化中,出现的一种数学思想是图论思想,它需要用一条连续的曲线画出一定的图形,而手指却不离开沙地.在西方数学中,这种思想曾被欧拉于 1736 年首次研究(见第 14 章).布松的孩子们在 1905 年第一次把他们的图表展现给欧洲的一位人类学家.很明显,他们不仅清楚保证曲线能被连续地画出的条件,而且知道使这条曲线能被迅速画出的过程.对于首魁人来说,画这种图不仅是孩子们的一种游戏,而且还是老年人讲故事的习俗的一部分.作为其故事的一部分,点被用来表示人,而那些相当复杂的曲线则要求包含图形中的一些点,同时要排除图形中的另外一些点. 事实上, 绘图的过程就是建立一个由点组成的矩形表格, 然后把曲线插入其中(图 I.6).如果不专门研究这些图表,就不容易确定哪些点应该在曲线里面,哪些在外面.但首魁人所遵循的详细的绘图规则使他们能迅速地一笔画出这些曲

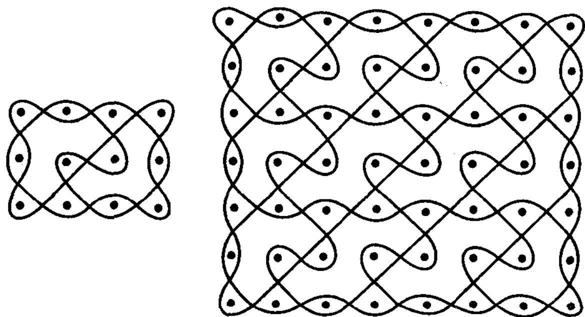


图 I.6　安哥拉首魁(Tshokwe)的条形图例.

线.[10]

另一个出现在许多非洲文化中的数学思想是几何图案,这被用在织布和金属品的装饰方面.在整个非洲有很多条形图案的例子,共含有7种常见的条形图案,以及其它17种平面图案.[11] 事实上,在扎伊尔巴库巴(Bakuba)人的布料上可以见到所有这7种条形图案和至少12种平面图案.尼日利亚班嫩(Benin)的艺术家们用所有这些条形图案和一些平面图案来装饰他们的青铜器(图I.7).

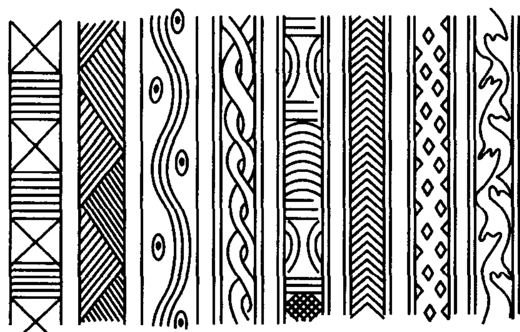


图 I.7　尼日利亚班嫩(Benin)的条形图例.

非洲也有数学游戏和谜题.比如,整个非洲都会玩一种被分别称为wari,omweso或mankala的拼板游戏,它对于培养孩子们的计算和策划能力很有用.类似地,一个熟知的谜题也出现在非洲的几种文化里,该故事讲述一个人准备把三个对象 A,B,C 运送到河对面,每次只能带一个,并且不能把 A 或 C 单独与 B 放在一起.在喀麦隆的巴米莱克(Bamileke),这三个东西分别指一只老虎,一只绵羊和一大捆芦苇.在其它地方则发现了一个不同的问题,那个人每次可以携带两个不同的东西.这些地方有阿尔及利亚(这三样东西是一只胡狼,一只山羊和一捆干草);利比里亚(一只印度豹,一只家禽和一些大米);桑给巴尔(一只豹子,一只山羊和一些树叶).这使人想起,这个谜题也出现在18世纪欧洲阿尔昆(Alcuin)的《命题》(Propositiones)中.

在南太平洋,我们也能在文努突(Vanuatu)共和国(澳大利亚东北1200英里处的一个岛国)的玛勒库拉(Malekula)发现在沙地上连续画出图形的思想.在玛勒库拉,这种思想潜伏于其宗教生活中.事实上,在那里,人们把准确地画出这些图形暗喻为寻找通往天堂之路.玛勒库拉人设计了一些标准的算法,通过对一些基本图形的对称操作来完成复杂图形的绘画.因此,可以用现代群论的语言来分析玛勒库拉人的图形.

群论在分析玛勒库拉人的亲属关系时,也是很方便的.事实上,老人们就是用图形向一位人类学家解释这些关系的.这些图形可以很容易转变成群的格式.其基本的思想是把群体分成六部分,一部分中的男人只能和一不同部分的女人结婚,而他们的孩子则又属于另一部分.如果一个男性属于我们称之为 e(身份)的部分,那么他父母将分别属于 f 和 m 部分,从而他奶奶属于 mf 部分,而其外公将属于 fm 部分.这样,亲属规则就是:m 和 f 的所有可能的"乘积关系"形成了一个6阶二面体群,即元素 m、f 通过关系 $m^3 = e$,$f^2 = e$ 和 $(mf)(mf) = e$ 所生成的6个元素组成的群.如果 B 属于 A 的奶奶的那一部分,或者,同样地,如果 A 属于 B 的奶奶的那一部分,那么婚姻只能在 A、B 之间发生.在澳大利亚北部的瓦尔皮里(Warlpiri)出现了一个相似的亲属关系的8阶群结构.[12]

在南太平洋的另一个地方马绍尔(Marshall)群岛,由棍棒组成的航海图是航海传统的一个要素.有些棍棒航海图的结构包括一些理想化了的形状,以训练海员掌握海浪运动的原理,尤其是海浪与海岸的相互作用原理.另一些模型在本质上则是整个马绍尔海域或其局部的地图.很明显,在这两种情况下,航海者已经构造出了数学模型,用以表示从一个岛航行到另一个岛所必须了解的关

于风与水的复杂相互作用的最重要的因素.这些模型一代代地传了下来,所以新一代的航海者就具备了必要的知识(图 I.8).[13]

在南太平洋,数学也用于满足娱乐的需求.比如,新西兰的毛里人玩一种叫做 mu torere 的两人游戏,每方可以使用 4 个标记杆,它们被放在八角星形状的游戏板上.对该游戏的彻底分析涉及组合学中的许多公式.两个水平相当的人玩该游戏时,需要对多种机会做仔细地思考与选择,以免自己出现漏洞并且令对手被困.



图 I.8 马绍尔群岛的棒式航海图.

对各种民族数学的简短介绍,使我们明白,两种数学的核心思想,即逻辑思维和模型分析,在世界各地都出现了.尽管大部分社会和中国、印度、伊斯兰及欧洲那些具有文字文明的社会不一样,没有正式的数学,但是,数学,不管在过去还是现在,都是世界各地人民生存的一种力量.

习 题

1. 解释那些给定一个日期和一段间隔天数之后,确定新的玛雅日期的公式的有效性.
2. 给定玛雅日期(8,10,193),确定在 0,2,3,5,10 天之后的新玛雅日期.
3. 找出一种算法来确定两个玛雅日期(t_0, u_0, y_0)和(t_1, u_1, y_1)之间的最短间隔天数.可以先忽略 365 天的历年,在 260 天的历年中确定两个日期(t_0, u_0)和(t_1, u_1)之间的最短天数.此题可查阅注释 4 中提到的 Closs 和 Lounsburg 的论文.
4. 说明两个玛雅日期(8,20,13)和(6,18,191)之间的最短天数是 1,8,15,18(= 10398).另外,由于这两个日期是一位玛雅国王 Pacal 的生卒日,且从别处资料可知他逝世时年龄在 60 岁和 100 岁之间,试确定 Pacal 共活了多少天及多少年.(注意,1 历日圈 = 18980 天 = 52 历年.)
5. 试完成玛勒库拉社会中关于其亲属关系的那个群表.对于每一部分的一位女性,确定其丈夫、母亲、父亲和孩子所在的部分.
6. 写一篇关于 7 种条形图案和 17 种平面图案的报告.从围墙纸或棉织品中找到所有种类的案例.参阅 D.K. Washburn 和 D.W.Crowe, *Symmetries of Culture*(Seattle: University of Washington Press,1988).
7. 研究 *mankala* 拼板游戏,并为孩子们设计一节课用这个游戏讲解各种算术概念.参阅 Laurence Russ, *Mancala Games*(Algonac, Mich. : Reference Publications, 1984),H.J.R.Murray, *A History of Board Games Other than Chess*(Oxford: Clarendon Press,1952),M.B.Nsimbi, *Omweso, a Game People Play in Uganda*(Los Angeles: African Studies Center, UCLA,1968).
8. 阅读文献 Anna Sofaer, Rolf M.Sinclair, Joey B.Donahue,"Solar and Lunar Orientations Of the Major Architecture of the Chaco Culture of New Mexico,"in *Proceedings of the Colloquio Internazionale Archeologia e Astronomia*(Venice,1990)以及文献中提到的其它参考书.这些文章能使你确信 Anasazi 人在方位确定和建筑设计中曾利用过数学工具吗?还有哪些其它的证据值得研究?
9. 阅读注释 6 中所引用著作(*Code of the Quipu*)的第 6 章,该章用某种树图的数学结构解释绳结语.为了在各种用途中创造出恰当的绳结语,印卡绳结语设计者必须作出何种树图分析?

文献和注解

1. 论述现代科学为什么出现在西方而非伊斯兰或中国的著作有:

Toby E. Huff, *The Rise of Early Modern Science: Islam, China, and the West*(Cambridge:Cambridge University Press,

1993), H. Floris Cohen, *The Scientific Revolution: A Historiographical Inquiry*(Chicago: University of Chicago Press, 1994), F. Jamil Ragep and Sally P. Ragep, eds., *Tradition, Transmission, Transformation*(Leinden: E. J. Brill, 1996).

2. 引自 H. Floris Cohen, *The Scientific Revolution*, P. 367.

3. 此图片源于 Persis B. Clarkson, "Classic Maya Pictorial Ceramics: A Survey of Content and Theme," in Raymond Sidrys, ed., *Papers on the Economy and Architecture of the Ancient Maya*(Los Angeles: Institute of Archaeology, UCLA., 1978), 86 – 141. Clarkson 鉴定此人为女性抄写员, Michael Closs 鉴定她为数学家, 因为数符饰带在她腋下.

4. 更详细的玛雅数学技能参见 Michael Closs, "Tte Mathematical Notation of the Ancient maya," in *Native American Mathematics*, Michael P. Closs, ed. (Austin: University of Texas Press, 1986), 291 – 369; Floyd G. Lounsbury, "Maya Numeration, Computation, and Calendrical Astronomy," in *Dictionary of Scientific Biography*(New York: Scribners, 1978), Vol. 15, 759 – 818.

5. 出处: Marcia Ascher. 同意使用.

6. 见 Marcia Ascher 和 Robert Ascher, *Code of the Quipu: A study in Media, Mathematics, and Culture*(Ann Arbor: University of Michigan Press, 1980). 该书最近再版为 *Mathematics of the Incas: Code of the Quipu*(New York: Dover Publications, 1997). 它不仅提供了绳结语制作技能的数学分析, 还有一些练习帮助学生研究相关的数学思想.

7. Claudia Zaslavsky, *Africa Counts: Number and Pattern in African Culture* (Boston: Prindle, Weber and Schmidt, 1973) 是最完整的关于非洲数学的著作. 含有世界上其它地区的一部近期著作是 Marcia Aster, *Ethnomathematics: A Multicultural View of Mathematical Ideas*(Pacific Grove, Ca.: Brooks/Cole, 1991). 这两部著作不仅提供了非常详细的数学思想, 还含有进一步阅读的大量文献. *For the Learning of Mathematics* 14(2)(1994) 是一期这方面的专刊. The Newsletter of the International Study Group on Ethnomathematics 提供这个领域的当前信息.

8. 关于阿纳萨支的详细资料, 查阅 William Ferguson and Arthur Rohn, *Anasazi Ruins of the Southwest in Color*(Albuquerque: The University of New Mexico Press, 1986). 关于阿纳萨支及其它北美印地安人的天文学, 见 Ray A. Williamson, *Living the Sky: The Cosmos of the American Indian* (Norman: University of Oklahoma Press, 1987), E. C. Krupp, ed., *In Search of Ancient Astronomies*(New York: Mc Graw – Hill, 1978).

9. 关于莫桑比克及其它非洲南部文化中的数学思想, 详见 Paulus Gerdes, "Geometrical-educational Explorations Inspired by African Cultural Activities"(Washington: MAA, 1998).

10. 详细的绘图规划, 见注释 9 中的书, 也见 Paulus Gerdes, "On Mathematical Elements in the Tchokwe 'Sona' Tradition," *For the Larning of Mathematics* 10 (1990), 31 – 34, 图 I.6 取自其中. 也可参阅 Ascher, *Ethnomathematiss*, 第 2 章.

11. 见 Zaslavsky, *Africa Counts*, 第 14 章. 这一章由 D. W. Crowe 撰写, "Geometric Symmetrics in African Art," 图 I.7 取自其中, 同意使用.

12. 见 Ascher, *Ethnomathematics*, 第 3 章.

13. 关于马绍尔群岛的棒式航海图, 详见 Marcia Ascher, "Models and Maps from the Marshall Islands: A Case in Ethnomathematics," *Historia Mathematica* 22(1995), 347 – 370.

第三篇

早期近代数学

1400—1700

第9章　文艺复兴时期的代数

> 对于三次方程，其中含有三次项，一次项及常数项，无法给出一般的解法.…… 当然，并不排除对某些具体的方程，通过试验，有时是可解的.
>
> ——引自帕乔利(Luca Pacioli)1494年的《算术、几何、比与比例集成》(Summa de Arithmetica, Geometria, Proportioni et Proportionalita)[1]

关于解三次方程的代数法则，卡尔达诺(Girolamo Cardano)在《大术》(Ars Magna)第11章中，给出了他的说明，"在三十年前[即，1515年]，波伦亚的费罗(Scipio Ferro)就发现了这个[解含有三次项，一次项及常数项的方程的]法则，并传给了威尼斯的菲奥尔(Antonio Maria Fior)，菲奥尔向布雷西亚的塔尔塔利亚(Niccolo Tartaglia)发起的挑战，促成塔尔塔利亚也发现了它.在我的恳求下，塔尔塔利亚向我吐露了解法，但拒绝了证法.由于这样的帮助，我终于找到了在各种情况下的证明过程.这是非常困难的."[2]

14世纪欧洲经济的巨大变化对数学具有一定的影响.随后两个世纪的文化运动，即文艺复兴对数学产生了更大的冲击，特别在意大利，因此，我们先谈那里的情况.

中世纪的意大利商人可称之为冒险资本家.他们航行到远东，买回所想要的货物，然后返回意大利卖掉货物来谋利.这些商人所需数学只不过是确定每次航行的收入和支出.然而，到14世纪早期，源于十字军东征的需要所引发的一场贸易革命大大地改变了这种系统.新的造船技术及航道上较大的安全性使文艺复兴中坐着不用动的商人得以替换中世纪时期的必须亲自航行的商人.这些"新的商人"可以留在意大利家中而雇用别人到处去航行，由那些人像代理商一样负责航程及具体的买卖.这种商业便导致国际贸易公司在意大利主要城市的诞生，这些公司相比他们的前辈需要更为复杂的数学.他们必须处理信用证、汇票、期票及利息计算.作为记录各种交易的方式，复式簿记开始出现.商业不再是一次性的货物交易，而是包含着在各个港口的多次装运，由一系列同时发生的货物流动所组成.基本上是易货方式的中世纪经济，逐渐便被今天的货币经济所替代.

意大利商人需要一种新的数学能力来适应这种新的经济形式，但他们所需要的并不是大学里

所学的那些四艺中的数学.他们需要解决实际问题及计算方面的新工具.为了满足这种需求,在 14 世纪早期的意大利,一种新的"职业"数学家阶层,算图学家(maestri d'abbaco),便应运而生,他们在专为此而建的新校中,给那些商人子弟编写课本并传授必要的数学知识.

本章第 1 节将讨论意大利算图学家的数学知识,特别是他们的代数.由于这场贸易革命不久便伸展到欧洲的其它部分,故第 2 节将讨论在 15 世纪后期和 16 世纪早期法国、德国、英国和葡萄牙的代数.但是,由于这个时期代数上的新发现主要出现在意大利,另外,也是为了答复本章开始所引用的帕乔利在 1494 年的断言,认为一般的三次方程是代数不可解的,我们将在第 3 节返回意大利,介绍关于其解法的最终发现的那段奇妙的故事,其中涉及费罗、塔尔塔利亚、卡尔达诺和邦贝利(Rafael Bombelli)的贡献.

所有上述代数学家的工作都是基于 12 世纪时最早译成拉丁文的伊斯兰代数.但到 16 世纪中期时,所有幸存的古希腊数学著作,从藏于君士坦丁堡的希腊原稿中被重新译成拉丁文,欧洲数学家已能看到了.因此,本章最后一节将专述韦达(Francois Viete)和斯蒂文(Simon Stevin)的贡献,韦达根据他对希腊数学的理解全面改进了代数学的研究,而斯蒂文彻底排除了亚里士多德对数与量的区分,有效地提出了我们当今的"数"的概念.

9.1　意大利的算图学家

14 世纪的意大利算图学家在给商人传授"新的"印度 —— 阿拉伯十进制位值系统及其算法规则时,奉行的是工具主义的哲学.象通常一样,当新的体系替代传统体系时,都会遇到巨大阻力.许多年中,帐簿仍然保持着罗马数字的形式.在支票上用文字写出帐目源于此时.人们认为,在记录大笔交易时仅用印度 —— 阿拉伯数字,容易被篡改.然而,新体系的优点最终战胜了开始时的犹豫.旧算板体系不仅需要一个算板,而且需要一袋筹码随身带着,而新体系只要有笔和纸就能到处使用.另外,在算板上操作、计算最终答案时,前面的步骤会被毁掉.而新体系的每一步都可留下以供复查.(当然,如果那里没有不久前刚出现的便宜而稳定的纸张供应,这些优点就无从谈起.)算图学家们用新算法教育着整个意大利的中产阶级后代,这些方法不久便传遍欧洲大陆.

除印度 —— 阿拉伯数系的算法之外,算图学家利用算术及伊斯兰代数为工具传授学生解题办法.算图学家所编课本,其中有几百种不同版本仍保存着,一般是问题及其解法的大汇编.[3] 其中不仅包含纯商业问题,这种题目学生将来进入父辈公司时会用到,而且还包含大量娱乐性问题,这类题目在现代的代数基础书中仍有典型性.有时还有一些几何题以及数论、历日及天文方面的题.课本中的解法非常详细,每一步都有充分的描述.但一般不解释各个步骤的原因,及某一具体方法的局限性.课本也不告诉读者当某一方法失效时该怎么办.我们也不知道教师在课堂上是否讨论这些情况.总之,这些算图课本不仅供课堂使用,而且也作为商人的参考手册.某一具体问题的解法很容易从书中找到,并可毫无困难地仿效,而不必知道其解法背后的理论知识.

书中大多数问题都可用古代的三分律或试位法解之,下面便是其中几例:

> 在[意]卢卡,1 弗罗林值 5 里拉,12 索尔多,6 迪拉里尼斯.问 13 索尔多,9 迪拉里尼斯值多少弗罗林?(1 里拉 = 20 索尔多,1 里拉 = 12 迪拉里尼斯)
>
> 1 里拉每月的利息是 3 迪拉里尼斯.问 60 里拉 8 个月挣多少利息?(这是一个单利问题.复利问题书中也有,每期的时间一般为一年.)
>
> 一块田 150 英尺长,一只狗和一只兔分站一端.狗一次跳 9 英尺,而兔一次跳 7 英尺.

问在多少尺、经过几次跳跃狗便抓住兔子?

虽然这些课本完全是实用性的,但它们对数学的发展曾有显著的影响,因为它们给意大利商人阶层逐步灌输了一种数字能力,而没有这种能力进一步的发展不可能出现.更进一步讲,其中某些课本作为课程的基本部分,给中产阶级带来了对伊斯兰代数的研究.在14世纪和15世纪中,算图学家曾在几个方面拓展了伊斯兰方法.特别地,他们引进了缩略词和符号表示,发现了处理复杂代数问题的新方法,把代数法则延展到了高于二次的方程中.然而,相比引入几种新技巧来说,更重要的是对如何把代数用于解决实际问题的全面讲授.通过对算图学家课本的研究,随着代数能力的日益增强,欧洲学者自然会转向应用这些技能解决更理论化的问题,即许多古典希腊数学原著被重新发现所引发的问题.代数和希腊几何学的结合在17世纪引发了新的分析技能,而这种分析技能后来成为现代数学的基础.

9.1.1 代数的符号表示和技能

伊斯兰代数完全是用言辞表达的.没有符号表示未知数及其幂次,也没有符号表示这些量的运算.一切都写成文字形式.在早期算图学家及更早的斐波那契的著作中,一般也如此.然而,在15世纪早期,一些算图学家开始用缩略词代替未知数.例如,有些作者分别用 *c*, *ce*, *cu* 和 *R* 代替词汇 cosa(物),censo(平方),cubo(立方)和 radice(根).这些缩略词的组合用以表示更高的幂次.因此,*ce di ce* 或 *ce ce* 表示四次幂;*ce cu* 或 *cu ce* 表示五次幂(x^2x^3);*cu cu* 表示六次幂(x^3x^3).然则,到15世纪末期,对高次幂的命名系统有了变化,作者用 *ce cu* 表示六次幂($(x^3)^2$),用 *cu cu* 表示九次方($(x^3)^3$),十五次方被表示为 *p.r.*,七次方为 *s.r.*

在15世纪末,帕乔利引入 $\bar{p}$ 和 $\bar{m}$ 表示加和减(più 和 meno).(这种缩写法可能来自于过去给单词上面加一横表示该单词在原稿中被遗忘的习惯.)然而,和其它革新一样,并未引起大家的立即响应.这个变化是一个缓漫的过程.新的符号在15世纪和16世纪得以使用,但真正意义上的现代代数符号表示的形成则是17世纪中期的事.

既使缺少一定的符号表示,意大利算图学家和他们的伊斯兰先辈一样,也能有效地进行代数运算.例如,格拉尔迪(Paolo Gerardi)在他1328年的《Libro di ragioni》中,给出了分数 $100/x$ 和 $100/(x+5)$ 的加法法则:

> 让100对应一物[x],再让100对应一物加5.然后如下图所示交叉相乘,…,100乘以一物得100物,100乘以一物加5得100物加500.现在,将两数相加得200物加500.然后一物乘以一物加5得1平方幂加5物.最后,用1平方幂加5物除200物加500得[$(200x+500)/(x^2+5x)$].[4]

100 ╳ 1 cos *a*
100 ╳ 1 cos *a piu* 5

类似地,正负号法则也被写成文字形式,甚至被验证,下面是14世纪晚期的一本无名作者的手稿:

> 负号乘以负号得正号.可以这样证明:我们知道,$3\frac{3}{4}$ 的自乘等于4减1/4的自乘.$3\frac{3}{4}$ 乘以 $3\frac{3}{4}$ 得 $14\frac{1}{16}$.而对4减1/4乘以4减1/4来说,由于4乘4得16,两个4乘以负1/4得负2,从16中减去2得14.现在,让负1/4乘以负1/4得1/16.故与乘法一样,得到了相同的结果.[5]

一般来说,采用上面所给出的未知数的幂的缩略词,算图手稿都会列举出单项式的乘积和商.唯有一本15世纪的手稿,在给出未知数的前九个幂次的命称后,还准确地阐述了指数法则:

> 两个幂相乘,… 可把系数相乘后,再把幂的次数相加.…… 两个幂相除,可把系数相除后,再把幂的次数相减.[6]

佛罗伦萨的S.Trinita修道院的算图学家安托尼奥(Antonio de'Mazzinghi)(1353—1383),其生平不详.他的代数问题在几本15世纪的手稿中均可找到.善于在解复杂的代数问题时,设计精巧的代数方法.特别是,在许多问题中对两个未知量采用两个不同的命称.例如,对于下面的问题:"求两个数,使其相乘为8且平方和为27."[7] 这位算图学家在解题之前,先设第一个数为 un cosa meno la radice d'alchuna quantità(一物减某一量的平方根),而设第二个数为 una cosa più la radice d'alchuna quantità(一物加某一量的平方根).其中cosa和quantità相当于我们今天的符号 x 和 y,即,第一个数等于 $x-\sqrt{y}$,第二个数为 $x+\sqrt{y}$.

9.1.2 高次方程

意大利算图学家的第三个主要创新是,把伊斯兰解二次方程的技巧扩展到高次方程.一般来说,算图学家开始论述代数学时,总是先列出花拉子米关于线性和二次方程的六种类型并展现出其解法.但是,比萨的达尔迪(Maestro Dardi),在1344年的一书中,将其拓展为包括4次方程在内的198种类型的方程,其中有些涉及到根式.[8] 虽然对每一情况,达尔迪都会给出新解法,既有数字例子又有解法诀窍,但其绝大多数都可适当简化为某一标准形式.例如,他指出方程 $ax^4=bx^3+cx^2$ 的解如下:

$$x=\sqrt{\left(\frac{b}{2a}\right)^2+\frac{c}{a}}+\frac{b}{2a},$$

即,它与标准方程 $ax^2=bx+c$ 同解.(注:0从不认为是一个解.)类似地,方程 $n=ax^3+\sqrt{bx^3}$ 是可解的,因为可把它化为关于 $\sqrt{x^3}$ 的二次方程.

更为有趣的是其中的四个有关三次和四次方程的例子.达尔迪的三次方程是 $x^3+60x^2+1200x=4000$.他的解法是,1200除以60得20,将这个结果三次方得8000,再加4000得12000,然后取立方根得 $\sqrt[3]{12000}$,最后减去1200除以60的商.达尔迪的答案是正确的,即 $x=\sqrt[3]{12000}-20$.如果用现代符号表示这个方程并给出达尔迪的解法,便得方程 $x^3+bx^2+cx=d$ 的解为:

$$x=\sqrt[3]{\left(\frac{c}{b}\right)^3+d}-\frac{c}{b}.$$

很明显,在一般情况下这个解是错的,甚至达尔迪也是承认的.那么,达尔迪是怎样对他的特殊情况算出正确解的?我们可以通过分析下面的复利问题来作出回答:某人租借给别人100里拉,三年后收回本利和总额为150里拉,其利息每年按复利计.利率是多少?达尔迪假设1里拉每月的利率是 x 迪拉里尼斯.那么,1里拉每年的利息是 $12x$ 迪拉里尼斯,或 $(1/20)x$ 里拉.所以,一年后的本利和是 $100(1+x/20)$,三年后为 $100(1+x/20)^3$.因此,达尔迪的方程是:

$$100\left(1+\frac{x}{20}\right)^3=150 \quad 或 \quad 100+15x+\frac{3}{4}x^2+\frac{1}{80}x^3=150,$$

或最终为

$$x^3+60x^2+1200x=4000.$$

因为方程的左边来源于一个立方,通过加上一个恰当的常数,它能被恢复成完全立方和的形式.一

般来说,因为$(x+r)^3 = x^3+3rx^2+3r^2x+r^3$,要把$x^3+bx^2+cx$写成完全立方和的形式,必须找到$r$满足$3r=b$及$3r^2=c$,而只有当$b^2=3c$时条件才成立.在达尔迪的例子中,由于$b=60$且$c=1200$,所以条件成立且$r=c/b=20$.

达尔迪对于特殊的四次方程也给出了类似的解法,而弗兰切斯卡(Piero della Francesca,1420—1492),一位著名的画家和算图学家,在他的书《Trattato d'abaco》中更是将这种解法推广到五次和六次方程.但都没有确切说明,解法仅适用于可化简为$h(1+x)^n=k$的情形,其中$n=4,5,6$.另有一本这个时期的(无名)手稿,声称方程$x^3+px^2=q$可被解,只要设$x=y-\frac{p}{3}$即可,其中y是$y^3=3\left(\frac{p}{3}\right)^2y+\left[q-2\left(\frac{p}{3}\right)^3\right]$的一个解.这形式上正确的,但作者仅把一个三次方程换成了另一个三次方程.在他的数字例子中,他用试验来解新方程,但原方程也可用试验法来做.不管怎么说,虽然算图学家没有给出三次方程的完整解答,但和他们的伊斯兰先辈一样,他们在拼搏着并有部分收获.

帕乔利(1445—1517),后期的算图学家之一,曾在1470年代成为[天主教]方济会的修道士,并在其晚年在意大利各地讲授数学.他是一位著名的教师,有一幅由巴尔巴雷创作的他的油画现挂在那不勒斯博物馆,画中他正在给一青年讲解几何学(图9.1).教学之余,他给学生写了三本不同的算图课本.他对当时的教学现状不满.问题之一是缺乏教学材料,因此,他大约用了20年收集材料并在1494年完成了当时内容最全面的数学课本,该书也是最早印刷的数学书之一.这本《算术、几何、比与比例集成》共有600页,用的是托斯卡纳方言而非拉丁文.它不仅包含实用算书,而且包含代数、复式簿记,及实用几何.该书没有多少独特的创见.事实上,大量的代数问题直接取材于弗兰切斯卡的著作,而实用几何与斐波那契的著作相似.然而,它的综合性以及印刷早的事实使其流传很广,影响甚大.它是16世纪意大利数学家的必读物,也是这些人后来拓展代数领域的共同基础.然而,在讨论这些进展之前,我们先看一下欧洲其它地区的同期情况.现代代数不仅仅是源于意大利.



图9.1 意大利邮票上的帕乔利.

9.2 法国、德国、英国和葡萄牙的代数

14、15世纪的意大利,主要由于商业经济的需求,代数学得以发展.稍后,欧洲其它地区的经济也在变.也出现一些数学课本来满足社会的这种新需要.本节将讨论法国许凯(Nicolas Chuquet)的著作,还有德国的鲁多尔夫(Christoff Rudolff)、施蒂费尔(Michael Stifel)和朔伊贝尔(Johannes Scheubel),英国的雷科德(Robert Recorde),和葡萄牙的努涅斯(Pedro Nunes).他们的著作在代数方面非常类似,也类似于15世纪意大利的代数,故可以看出,这些数学家对欧洲其它地区的同期著作都有所了解,虽然对他人著作的直接照搬甚少或完全没有.显然,到15世纪时,伊斯兰代数知识在欧洲已广泛传播.他们都想利用这种材料及欧洲其它地区的代数著作写出新书,以使其适合自己国家的环境,并介绍一些他们自己的新思想.到16世纪后期,由于印刷术的普及,新思想得以更快地传遍欧洲大陆,其中最重要的东西便被吸收到一种新的欧洲代数中去.

9.2.1 法国:许凯

许凯(卒于1487)是一位法国医生,临去世前几年在里昂写过一部数学著作.15世纪晚期的里昂是一个繁荣的社会,和意大利城市一样,日益需要实用数学.也许是为了满足这种需求,许凯在

1484 年撰写了《算术三篇》(Triparty),该书分为三部分论述算术和代数,其后有三个附录,涉及多方面的问题,这些问题均可用《算术三篇》中建立的法则解决.[9] 这些附录中的题目类似于意大利算图学家的著作,但《算术三篇》本身首先是一部数学理论著作.其中大多数内容虽然是伊斯兰代数学家及比萨的斐波那契所熟知的.然而,由于它是 15 世纪法国的第一本系统的代数学,我们将分析其中的一些重要思想.

《算术三篇》的第一部分论述算术.像意大利的著作一样,开始先叙述印度 —— 阿拉伯位值系统及有关整数和分数的算术基本运算的各种法则.许凯关于分数的法则之一是"如何根据需要,在两个相邻数之间找到一系列的中间数".[10] 他的想法是,要在两个分数之间找一个分数,只需简单地把分子、分母分别相加即可.因此在 1/2 和 1/3 之间是 2/5,在 1/2 和 2/5 之间是 3/7.许凯没有证明这个法则的正确性,但他曾用其求多项式的根.例如,要求 $x^2 + x = 39\frac{13}{81}$ 的根,许凯先指出 5 太小而 6 太大,然后通过检查 $5\frac{1}{2}$,$5\frac{2}{3}$,$5\frac{3}{4}$,及 $5\frac{4}{5}$ 得知,根必须在最后两个值之间.对分数部分应用他的法则,他下一步检查 $5\frac{7}{9}$,恰好便得正确答案.

在《算术三篇》的第二部分,许凯用这个法则求数的平方根,该数不是完全平方数.首先指出,对于 6 的平方根来说,2 是太小而 3 太大,然后他便开始逼近,指出 $2\frac{1}{3}$ 是太小而 $2\frac{1}{2}$ 太大.他随后的几步逼近依次是 $2\frac{2}{5}$,$2\frac{3}{7}$,$2\frac{4}{9}$,$2\frac{5}{11}$,及 $2\frac{9}{20}$.他每一步都是先算出所选数的平方,再根据是否大于 6 来确定下一步在哪两个数之间应用他的法则求其中间数.他指出"可用这种方法进行下去,…… 直到非常接近 6,不管是大或小,只要能满足要求为止.人们应该清楚,逼近步骤越多,近似度越高.但永远不会得到精确值.我们所得 6 的平方根的最佳近似值是 $2\frac{89}{198}$,这个值自乘本身得 $6\frac{1}{39204}$".[11] 许凯明显知道 $\sqrt{6}$ 的无理性并发现了一种新的递归算法可将它计算到任意所想要的精度.因此可以说,在否定希腊人将离散与连续二分的合理性方面,许凯已有作所作为,而这种二分思想的最终消失大约发生在一个世纪之后.

在书中第二部分,许凯还用这种标准方法来计算较大整数的平方根和立方根.如果这种标准方法得不到精确根,还有另外一种选择,根本不用计算仅将结果记为 $R^2 6$ 或 $R^3 12$ 的形式,他的这个符号就是现在的 $\sqrt{6}$ 和 $\sqrt[3]{12}$.许凯还使用意大利的 $\bar{p}$ 和 $\bar{m}$ 表示加和减,但引入了一条下线指明分组情况.因此,现今的符号 $\sqrt{14 + \sqrt{180}}$,许凯当初写为 $R^2 14\bar{p}R^2 180$.在整个第二部分的其余部分,许凯便着手完全正确地使用这个符号以及含有正数和负数的四则运算法则,来处理根式运算.

《算术三篇》的第三部分主要讨论代数.许凯在此处论述了怎样处理多项式和怎样解各种类型的方程.在多项式的讨论中,他对未知数的幂引入了一种指数符号,使得运算比意大利的缩略词更为容易.因而,他用 12^2 表示现今的 $12x^2$,并且,在欧洲人的著作中首次实际引入负数,用 $\bar{m}12^{2\bar{m}}$ 表示现今的 $-12x^{-2}$.他甚至指出,如果仅涉及数字本身,其指数应表示为 0.可以用通常的标准法则,对这些含有指数的表达式进行加、减、乘、除四则运算,即使某一指数是负数也不受影响.因此,"欲用 8^3 乘以 $7^{1\bar{m}}$,首先需要用 8 乘以 7 得 56,然后指数相加,即 $3\bar{p}$ 加 $1\bar{m}$ 得 2.故这个乘积为 56^2,[两单项式相乘的]其它题目皆同此法."[12] 他不仅和一位意大利的同时代学者一样,给出了这个法则,而且他还验证了它.他将 2 的幂值(从 $1 = 2^0$ 到 $1\,048\,576 = 2^{20}$)及其对应的指数,列成两列,然后指出第一列的乘法对应于第二列的加法.例如,128 乘以 512 得 65 536,对应于 7 加 9 得 16.由于指数的加法法则对数字有效,他自然地将其推广到他的代数表达式.但是,虽然他理解负指数的意义,他的数字

表中却没有包含它们，事实上，不同于塞毛艾勒，他随后几乎没有使用它们.

许凯在解方程技巧上也有一些革新. 首先，他把花拉子米的法则推广到了本质上仍是二次型的一些任意次方程，因此，比意大利算图学家更深入. 例如，他给出方程 $cx^m = bx^{m+n} + x^{m+2n}$ 的解为

$$x = \sqrt[n]{\sqrt{\left(\frac{b}{2}\right)^2 + c} - \left(\frac{b}{2}\right)}.$$

其次，他指出含三个未知数的两个方程的方程组具有多重解. 为解方程组 $x + y = 3z, x + z = 5y$，他先把 x 取为 12，求得 $y = 3\frac{3}{7}, z = 5\frac{1}{7}$. 然后他把 y 取为 8，算出 $x = 28, z = 12$. 因此，他总结到，“相应于所取的数值，可确定出各种不同的答案.”[13] 最后，虽然他并未一贯坚持，许凯希望在某些场合应考虑方程的负解，这一观点在欧洲也属最早. 例如，他解方程 $\frac{5}{12}\left(20 - \frac{11}{20}x\right) = 10$，得 $x = -7\frac{3}{11}$. 详细检验后，他说答案是正确的. 然而，在有些问题中，他又否定负数解，认为“不可能”，而且从不认为 0 可为解.

《算术三篇》的三个附录含有数百问题，均用书中的技巧答之. 其大多数是商业问题，属于意大利算图著作中那些类型，也有一些是几何上的实用和理论问题.《算术三篇》可能打算作教材用，不一定用于大学，但很不幸，该书从未被印刷，今天所保存的也只是手稿形式. 其中有些内容曾在 1520 年，由洛歇(Estienne de La Roche) 编入他的书中，洛歇可能是许凯的一个学生. 但是，他们两人的书均未曾产生很大影响.

9.2.2 德国：鲁多尔夫、施蒂费尔和朔伊贝尔

德国代数最早出现在 15 世纪后期，引起代数发展的原因和早些时候的意大利相似. 事实上，很可能多数技能直接源于意大利. 连代数的德文名称(the art of the coss) 也显露着其意大利根源. Coss 是意大利词 cosa 的德文形式，cosa 的意思是“物”，在代数方程中通常表示未知数的名称. 16 世纪上半叶的两位最重要代数学家是鲁多尔夫(15、16 世纪之交) 和施蒂费尔(1487—1567).

鲁多尔夫于 1520 年代早期在维也纳写作他的著作《未知数》(Coss)[14]，该书于 1525 年在斯特拉斯堡出版，是第一本用德语写的综合性代数书. 和通常一样，著作开始是整数的位值制基础，并给出运算法则以及一个短乘法表. 在论述级数的一节中，鲁多尔夫列出一个和许凯一样的表，关于 2 的幂值及其相应的指数. 他也指出，幂的乘法对应于指数相加. 然后再次和许凯一样，把这种思想扩展到未知数的幂. 虽然鲁多尔夫没有他的法国前辈那样的指数符号，他却有一个关于这些幂的名称的缩略词系统，其中他的命名方案类似于意大利的倍增方案(补遗 9.1).

补遗 9.1 鲁多尔夫关于未知数幂的缩略词系统

dragma	φ	radix	ꝛ ↔ x
zensus	ʒ ↔ x^2	cubus	ꝯ ↔ x^3
zens de zens	ʒʒ ↔ x^4	sursolidum	ß ↔ x^5
zensicubus	ʒꝯ ↔ x^6	bissursolidum	bß ↔ x^7
zenszensdezens	ʒʒʒ ↔ x^8	cubus de cubo	cꝯ ↔ x^9

为了帮助读者了解这些术语，鲁多尔夫给出各种数字的幂作为例子. 然后，他展示了怎样对这些符号所表示的式子进行四则运算. 因为与许凯体系的情形不同，如何把这些符号相乘并不明显，

鲁多尔夫便做了一个乘法表供使用，例如，其中的 𝔵 乘以 𝔷 是 𝔠. 为了把事情简化，他还给他的符号添上数值编号. 因此，radix 被标为 1，zensus 为 2，cubus 为 3，等等. 他指出，把这些表达式相乘时，只要简单地把对应的数字相加，并查找正确的答案即可. 在这一节中鲁多尔夫还讨论了二项式，其中术语的联接用的是运算符，包括用现行的符号 + 和 − 来表示加和减，这在代数学中是第一次出现. 这些符号曾早些时候在鲁多尔夫的维也纳大学的老师，格拉马托伊斯(Henricus Grammateus)，1518 年的一本算术书中出现过. 甚至更早，这些符号还出现在维德曼(Johann Widman)1489 年的一书中，然而，它们在那里表示盈和亏，而不是这里的运算.

鲁多尔夫还在书中引入了现代符号$\sqrt{\ }$ 表示平方根. 他把这个符号略加修改来表示立方根及四次方根，但不同于现在的形式. 然而，他对根式运算作了详细论述，展示了在除法中怎样使用共轭数，另外还有怎样求根式的平方根，如$\sqrt{27+\sqrt{200}}$. 他还引入了用一个句点表示“相等”，如用 1 𝔵 .2 表示 $x=2$. 但是，他经常使用德文的相等(gleicd)这个字.

鲁多尔夫《未知数》的后半部分主要解代数方程. 并未用花拉子米关于方程的六种标准类型，鲁多尔夫采用了自己的八种分类. 各类方程的解法法则均用文字形式给出，并辅以例释. 虽然鲁多尔夫也处理过高于二次的方程，但与许凯一样，都能简化为二次方程解之或用简单的根解之. 例如，其中一类方程用现代形式可写为 $ax^n+bx^{n-1}=cx^{n-2}$. 所给的解是标准形式

$$x=\sqrt{\left(\frac{b}{2a}\right)^2+\frac{c}{a}}-\frac{b}{2a}.$$

所附的具体例子是 $3x^2+4x=20$ 和 $4x^7+8x^6=32x^5$，两个方程的解都是 $x=2$. 然而，像其他作者一样，鲁多尔夫没有考虑负根或把零作为根.

在呈现这些法则之后，鲁多尔夫给出数百个问题的例子，都可用法则解之. 大多数是涉及买卖、交易、资金和财产的商业问题，或娱乐问题，包括古老的百禽问题. 鲁多尔夫第一类方程 $ax^n=bx^{n-1}$ 的例子中的大多数问题，属于实用型的. 而需要二次方程的问题一般是编拟的，包括常见的“把 10 分为两部分，使之满足 ….” 在著作的最后，鲁多尔夫给出三个不可约三次方程及其解，而没有给出解法. 他轻松地声称，后辈学者会继续这门艺术并告知其解法. 令人费解的是最后一页的草图，该图把一个边长为 $3+\sqrt{2}$ 的立方体截成八个长方体. 是否鲁多尔夫想用这个图提示三次方程的解，那就不得而知了.

鲁多尔夫的著作曾由施蒂费尔在 1553 年再版发行，而施蒂费尔在九年前曾出版过一本他自己的著作《整数算术》(Arithmetica Integra).[15] 施蒂费尔在他的著作中使用了鲁多尔夫关于未知数幂的符号，但他更强调使用这些符号和整“指数”之间的对应. 在列出 2 的幂与其指数的对应表时，他更深入，其中包含负值 −1，−2，−3 对应于 1/2，1/4 和 1/8，但他可能不了解许凯关于负指数的类似工作.

人物小传

米歇尔 · 施蒂费尔(1487—1567)(Michael Stifel)

米歇尔 · 施蒂费尔于 1511 年成为牧师，是 Martin Luther 的早期追随者. 他在 1520 年代开始对《圣经》中的词语计算感兴趣，即用相关字母的数值解释词意. 通过用他的数值方法对某些《圣经》章节的解释，他最后确信世界的末日是 1533 年 10 月 18 日. 那天早晨他将全体教徒召集到教堂，但令他无比沮丧，什么也没有发生. 随后他被停职和软禁. 然而，在他放弃预言以及 Luther 的干预之下，他在 1535 年被任命到另一教区. 随后他在维滕贝格(Wittenberg)大学献身于数学研究并很快成为代数方法的专家，于 1544 年出版了《整数算术》，一年后还出版了另一本书《德国算术》(Deutsche Arithmetica). 然而在晚年，他又恢复了他的词语计算并写过两本相关的书.

虽然施蒂费尔和同时代学者一样,不接受方程的负根,但他首次把方程的三种标准形式压缩为一种形式 $x^2 = bx + c$,其中 b 和 c 或者同正或者奇偶性相反.而用文字所表示的解等价于

$$x = \frac{b}{2} \pm \sqrt{\left(\frac{b}{2}\right)^2 + c},$$

其中的负号仅用于 b 是正的,c 是负的.在这种情况下,只要 $(b/2)^2 + c > 0$,便有两个正解.把方程的三种情况归为一种似乎不算大的进步,但在16世纪的背景下,它是很重要的.下一步对数的概念的扩充,整整等待了两个世纪之久.

施蒂费尔书中还有关于二项式系数的“帕斯卡三角形”的描述,它是欧洲早期的版本之一(表9.1).(另一个早些的版本出现在阿皮安努斯(Peter Apianus)1527年的《算术》书的封页上,但阿皮安努斯在书中没有使用这个三角形.)施蒂费尔指出,他已发现了利用这些系数求[高次幂的]根的过程,不过太困难而无法书面描述.因此,虽然这些系数早在几个世纪前已被中国人和伊斯兰人用于相同的目的,但施蒂费尔这时才得知.

表9.1　帕斯卡三角形(施蒂费尔版本和朔伊贝尔版本)

施蒂费尔版本							
1							
2							
3	3						
4	6						
5	10	10					
6	15	20					
7	21	35	35				
8	28	56	70				
9	36	84	126	126			
10	45	120	210	252			
11	55	165	330	462	462		
12	66	220	495	792	924		
13	78	286	715	1 287	1 716	1 716	
14	91	364	1 001	2 002	3 003	3 432	
15	105	455	1 365	3 003	5 005	6 435	6 435
16	120	560	1 820	4 368	8 008	11 440	12 870
17	136	680	2 380	6 188	12 376	19 448	24 310

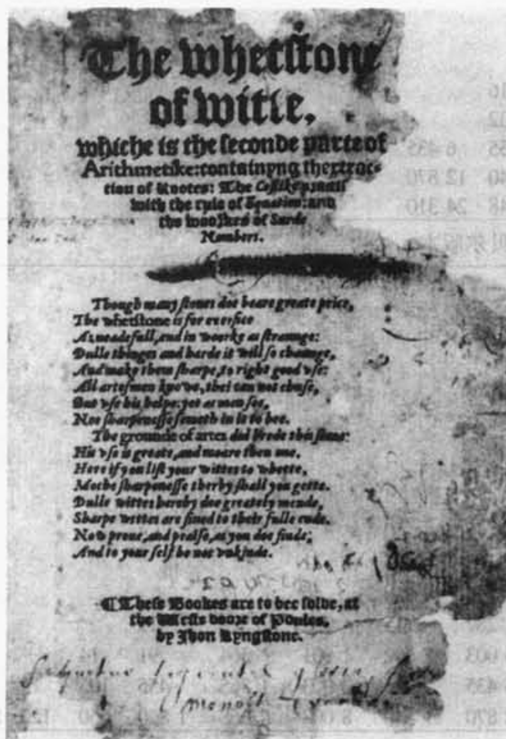
朔伊贝尔版本														
2														
3	3													
4	6	4												
5	10	10	5											
6	15	20	15	6										
7	21	35	35	21	7									
8	28	56	70	56	28	8								
9	36	84	126	126	84	36	9							
10	45	120	210	252	210	120	45	10						
11	55	165	330	462	462	330	165	55	11					
12	66	220	495	792	924	792	495	220	66	12				
13	78	286	715	1 287	1 716	1 716	1 287	715	285	78	13			
14	91	364	1 001	2 002	3 003	3 432	3 003	2 002	1 001	364	91	14		
15	105	455	1 365	3 003	5 005	6 435	6 435	5 005	3 003	1 365	455	105	15	
16	120	560	1 820	4 368	8 008	11 440	12 870	11 440	8 008	4 368	1 820	560	120	16

在随后几十年的德国人著作中，帕斯卡三角形也曾用于求根. 例如，朔伊贝尔(1494—1570) 在 1545 年的著作 De numeris et diversis rationibus 中，曾呈现出这个三角形及其制表指令. 朔伊贝尔的著作是拉丁文，显然用于不同的读者群. 特别地，他没有顾及材料的“实用性”，而是花费大量篇幅来探讨如何用帕斯卡三角的表值求高次幂的根的方法. 朔伊贝尔的 De numeris 不是代数课本，他在 1552 年曾在法国出版过一本代数课本(Algebrae compendiosa facilisque descriptio)，也是拉丁文. 这是除了洛歇版的许凯的《算术三篇》之外在那里出版的第一本代数著作.

9.2.3　英国：雷科德

斯蒂费尔的《整数算术》及其 1553 年对鲁多尔夫《未知数》的修订在德国是很重要的，甚至影响了下一世纪的著作者，并帮助德国中产阶级提高了数学认识. 它们也影响到英国，是英国最早的代数书《智力磨石》(The Whetstone of Witte, 1557) 的主要来源，该书作者是文艺复兴时期英国最早的数学书写作者，雷科德(1510—1558)(见图 9.2).

人物小传	罗伯特 · 雷科德(1510—1558)(Robert Recorde)
	雷科德 1531 年毕业于牛津大学且此后不久获得行医执照. 虽然在 1540 年代后期他曾可能在伦敦开业行医，但目前仅知他曾有过几种文职工作，俱不很成功. 另一方面，他却写出几本很成功的数学课本. 除了论述代数的《智力磨石》(1557)，还有关于算术的《技艺基础》(The Ground of Arts)(1543)，关于几何的《知识入门》(The Pathway to Knowledge)(1551)，以及关于天文的《知识宝库》(The Castle of Knowledge)(1556). 这些著作说明他对教学方法甚感兴趣. 他的著作体系是师徒对话形式，对具体技能的每一步骤都有详细解释.



The whetstone
of witte,
whiche is the seconde parte of
Arithmetike: containyng the extraction of Rootes: The Cossike practise
with the rule of Equation: and
the woorkes of Surde
Nombers.

Though many stones doe beare greate price,
The whetstone is for exersice
As neadefull, and in woorke as straunge:
Dulle thinges and harde it will so chaunge,
And make them sharpe, to right good use:
All artesmen knowe, thei can not chuse,
But use his helpe: yet as men see,
Noe sharpenesse semeth in it to bee.
The grounde of artes did brede this stone:
His use is greate, and moare then one.
Here if you list your wittes to whette,
Moche sharpenesse therby shall you gette.
Dulle wittes hereby doe greately mende,
Sharpe wittes are fined to their fulle ende.
Now proue, and praise, as you doe finde,
And to your self be not unkinde.

These Bookes are to bee solde, at
the Weste doore of Poules,
by Jhon Kyngstone.

图 9.2　雷科德《智力磨石》(1557) 的扉页.(出处：Smithsonian Institution Libraries, Photo No. 92 – 338.)

《智力磨石》在技能上几乎没有原创性,它完全基于德国人的资料,甚至表示未知数的幂也源于德国人的符号,但是,该书仍有几点有趣之处.首先,雷科德创造了相等的现代符号:“为了避免‘相等于’这个词的重复,正如我惯常的做法,我将引入一对等长的平行线,即═══,因为再找不到哪两种东西更能体现出相等的含意.”[16] 其次,他把德国人关于未知数的幂的符号体系修改和扩充到第80次幂的情形,在建立整数与这些符号的对应关系后,他指出符号相乘对应于相应的整数相加.事实上,他利用素数分解式,从各素数幂的符号出发,展示了如何建立任意次幂的符号.例如,素数幂中2次幂为𝔷,3次幂为𝔠,各个大于3的素数次幂为*ß,(其中*处表示一个素数的序的字母).5次幂为ß,7次幂为^{b}ß(表示3次幂后第2个素数幂),11次幂为^{c}ß(表示第3个).然后,比如说,9次幂写成𝔠𝔠(表示立方的立方),20次幂写成𝔷𝔷ß(5次幂的平方的平方),21次幂写成𝔠^{b}ß(7次幂的立方).最后,为了帮助学生记忆各种运算法则,他把它们编成诗歌的形式.他的诗句给出了单项式 ax^n 的乘法和除法程序,诗中的“量”表示指数 n,该诗说明了运算的记号的标准法则以及指数法则:

> 做乘者,
> 喜其得之更多,
> 做除者,
> 忧其剩之愈少.
> 其量亦有所变,
> 乘则须加除则须减.

9.2.4　葡萄牙:努涅斯

葡萄牙也需要数学,当地的航海家们在15世纪时就已经在吸取别国的知识.正是在这里,努涅斯(1502—1578)于1532年写出他的著作(Libro de Algebra).[17] 努涅斯曾受帕乔利著作的影响.他的符号明显来源于意大利的作者,他似乎并不了解同时代的德国学者的工作.因此,他用意大利的缩略词表示各种未知数的幂——co表示未知数cosa,ce表示平方censo,cu表示立方cubo——以及 $\bar{p}$ 和 $\bar{m}$ 表示加和减.他在书中论述了这些代数表达式的运算,解方程,根式及比例.他书中的几十个问题不同于前人,全部是抽象的.没有包含商业或娱乐问题.但他包含了一节内容论述代数技巧在几何上的应用.

人物小传

佩德罗·努涅斯(1502—1578)(Pedro Nunes)

佩德罗·努涅斯(图9.3)攻读于塞拉曼加(Salamanca)大学,但在里斯本大学获医学学位(1525).他曾在航海科学上做出几个贡献,并成为葡萄牙国王的首席宇宙学顾问和科英布拉(Coimbra)大学的数学教授,于1532年写出的代数著作(Libro de Algebra).他的许多学生后来成为宫廷高官.虽然努涅斯属于犹太血统,但他从未受到宗教法庭迫害,也许这与他的一位学生是葡萄牙宗教法庭庭长有关.努涅斯的代数课本最初是葡萄牙语,但他认为若改用西班牙语会更具影响力,他便在30年后将该书译成西班牙语且于1567年在荷兰出版.然而,他的天文著作主要用的是拉丁语.除他的科学著作之外,努涅斯还是一位有名的诗人.



图9.3　葡萄牙邮票上的努涅斯,注意背景的代数问题.

为了看出努涅斯著作的特点,我们来分析他是怎样解一个标准问题的,已知两数的积及其平方和,求这两数.假设积为 10,平方和为 30.努涅斯用了三种方法解题,以显示不同的代数技巧.我们将用现代符号,譬如,用 x 而不用努涅斯的 co.首先,设 x 是两数中较小者,他取 $10/x$ 为较大者,平方每一个,得方程

$$x^2 + \frac{100}{x^2} = 30.$$

用 x^2 乘以这个方程的两边,将其化为关于 x^2 的二次方程,并用花拉子米的公式之一解知 $x^2 = 15 \pm \sqrt{125}$.因此,他便得到这两数为 $\sqrt{15 - \sqrt{125}}$ 和 $\sqrt{15 + \sqrt{125}}$.对于第二种方法,他指出,两数不可能相等,然后,用 $15 - x$ 和 $15 + x$ 表示这两数各自的平方.用 $\sqrt{15 - x}$ 和 $\sqrt{15 + x}$ 表示两数本身,他由此导出方程 $\sqrt{15 - x}\sqrt{15 + x} = 10$,这容易化简出 $x = \sqrt{125}$.因此,解和前面一样.努涅斯的第三种解法用到恒等式 $(a + b)^2 = a^2 + 2ab + b^2$.两数和的平方是 50,故两数和是 $\sqrt{50}$.因此,这两数是 $(1/2)\sqrt{50} - x$ 和 $(1/2)\sqrt{50} + x$.将其相乘,便得方程

$$12\frac{1}{2} - x^2 = 10,$$

所以 $x = \sqrt{2\frac{1}{2}}$.那么,此时这两数为 $\sqrt{12\frac{1}{2}} - \sqrt{2\frac{1}{2}}$ 和 $\sqrt{12\frac{1}{2}} + \sqrt{2\frac{1}{2}}$.努涅斯现在的问题是如何说明这组数与前面两种方法得到的那组数相同.既使他清楚平方相等未必一定根就相等,他还是通过比较各自的平方来说明问题.虽然他不清楚怎样克服这个困难,但他还是相信解在本质上是相同的,因为两对数都满足原方程.

9.3　三次方程的求解

[意]帕乔利在 1494 年指出,对一般的三次方程还没有得到代数解,但整个 15 世纪和 16 世纪早期,许多数学家都在探索这个问题.最终,在 1500 年和 1515 年之间的某个时期,波伦亚大学的教授,费罗(1465—1526),发现了解三次方程 $x^3 + cx = d$ 的一种代数方法.通过我们对当时伊斯兰关于三次方程求解的研究可以知道,数学家一般不涉及负数,既使方程的系数也如此,因此,根据二次项、一次项和常数项的相关位置不同,共有 13 种类型的混合不可约三次方程.所以,只能说费罗通过对其中一类的求解,才刚开始了解三次方程的过程.

在当今学术界,教授们为了保证优先权,总是尽可能快地公布和发表他们的新成果,故可能感到奇怪,为什么费罗未曾发表,甚至公开宣布这个主要突破点.但是,16 世纪意大利的学术环境与现在完全不同.当时没有著作权.大学职位很不稳定,由校方评议会定期更换.教授使评议会相信他值得继续保持他的职位的手段之一是,赢取公开挑战.某一职位的两个竞争者需要互相解答对方的问题.除了大学职位本身,可观的奖金通常也依赖于这种挑战的结果.这样一来,如果某一教授对解决某些问题发现了新方法的话,保密便对他有利.他然后便可在定能胜出的领域向对方有把握地发问.

费罗在他去世前,曾将其解法透露给他的学生菲奥尔(16 世纪上半叶)以及他在波伦亚的继任者纳夫(Annibale della Nave,1500—1558).虽然这两个人都未曾将解法公开,但消息还是在意大利数学家中开始流传了,这个古老的问题已经被,或者将要被解答.事实上,另一位数学家,布雷西亚的塔尔塔利亚(1499—1557)曾自夸他也发现了一种三次方程 $x^3 + bx^2 = d$ 的解法.在 1535 年,菲奥尔向塔尔塔利亚发起公开挑战,希望凭借他对前一种情形的知识而取胜.他所提交的 30 个问题全

部属于那种类型的方程.例如,其中一个问题是,"一块蓝宝石卖了500金币,所得利润是其成本的立方根,求其利润是多少?"[$x^3 + x = 500$].[18]塔尔塔利亚,这位更优秀的数学家,经过对那种类型方程夜以继日的研究,终于像他后来记述的那样,在1535年2月12日发现了其解法.由于菲奥尔未能解出塔尔塔利亚的大多数问题,那些问题除了涉及后一类型的三次方程外还需要数学的其它知识,塔尔塔利亚便被宣告为胜者.在这种情况下,输者需要宴请胜者及其朋友,不过,塔尔塔利亚明智地谢绝了这种奖赏,仅接受了荣誉.

竞赛的消息及三次方程的新解法不久传到了米兰,其时,卡尔达诺(1501—1576)在学者皮亚蒂(Tommasso Piatti)遗嘱的赞助下,正在为贫民青年开设数学公开讲座.卡尔达诺写信给塔尔塔利亚请求告知其解法,以便能将该解法完全以塔尔塔利亚的名义包含在卡尔达诺那时正在写作的一本数学课本中.塔尔塔利亚最初予以拒绝,但在卡尔达诺多次恳求并承诺将把他及他关于火炮方面的发现引见到米兰宫廷之后,他最终于1539年早期来到米兰.在卡尔达诺发誓绝不发表塔尔塔利亚的发现之后——他计划在随后某个时期亲自发表其解法——塔尔塔利亚以诗歌的形式向卡尔达诺泄漏了三种不同形式的三次方程的秘密.下面是其中解释 $x^3 + cx = d$ 的一段诗文:

立方共诸物,　　和要写右边,
巧设两个数,　　差值同右和;
此法要牢记,　　再定两数积:
诸物三(分)之一,　还把立方计;
既知差与积,　　两数算容易,
复求立方根,　　相减题答毕.[19]

人物小传	格罗拉莫·卡尔达诺(1501—1576)(Gerolamo Cardano)
	卡尔达诺早年毕业于医学专业,但因他是私生子而被拒绝加入米兰医生协会.因此有几年时间他先在帕度亚(Padua)附近的一个小镇独自开业行医,后于1533年返回米兰一边接收一些临时病人,一边讲授数学并撰写算术课本.最终,他使米兰医生协会改变了原来的决定.卡尔达诺很快成为米兰最杰出的医生并闻名全欧.卡尔达诺最重要的病人是苏格兰大主教,Sohn Hamilton,他于1551年邀请卡尔达诺治疗他不断恶化的气喘发作.卡尔达诺对主教的症状和生活习惯经过一个月的观察,认为他对装着羽毛褥垫的床有严重的过敏反应.因此卡尔达诺建设将床上用品全部换成丝制和亚麻织物.主教的健康立即好转,他在余生对卡尔达诺甚为感激,不断给予经济和其它帮助.但卡尔达诺从苏格兰回国途径英格兰时,以占星术为年青的英王爱德华6世所算的命却不那么成功.他预言能够长寿,但不幸地,此后不久16岁的英王便去世.卡尔达诺的个人生活也充满悲剧,1546年妻子去世,1560年儿子因谋杀自己的妻子被处死刑.最后的打击发生在1570年,他被控信奉异教邪说而接受宗教法庭的审判.幸运地,判决还算宽大.卡尔达诺的最后几年在罗马渡过,在那里他写出自传 De Propria Vita.

9.3.1 卡尔达诺和《大术》

卡尔达诺保持了自己的承诺,在其后不久完成的一部数学书中并没有发表塔尔塔利亚的成果.事实上,他还寄给塔尔塔利亚一个副本以显示自己的良好信誉.随后,卡尔达诺便开始钻研这个问题,也许曾得到他的家仆和学生费拉里(Lodovico Ferrari,1522—1565)的帮助.经过几年的努力,他

终于对各种类型的三次方程都得到了解法及其证明.费拉里本人还解答了某些四次方程.而在此期间,塔尔塔利亚仍未发表三次方程的任何东西.卡尔达诺并不想违背自己的正式誓言,但他渴望该解法能被公众所知.针对费罗是原始发现者的传言,他和费拉里曾旅行到波伦亚拜访过纳夫.纳夫宽厚达理地允许他们查阅费罗当初的手稿,使他们得以核实费罗最早发现解法.卡尔达诺便不再觉得对塔尔塔利亚有义务.无论如何,他将发表的解法从根本上讲,是二十多年前费罗的发现,而费罗当时已亡故.因此在 1545 年,卡尔达诺出版了他最重要的数学著作《大术》(the Ars Magna),又名《论代数法则》,主要致力于三次和四次方程的求解(图 9.4).当然,当卡尔达诺著作出现时,塔尔塔利亚会狂怒.他感到自己的劳动成果已被别人骗取,既使卡尔达诺已经提到塔尔塔利亚是这个方法的最初发现者之一.塔尔塔利亚的抗议并未对他有任何帮助.为了恢复自己的威望,他进行了另一次公开竞赛,这次是和费拉里,但是他失败了.直到现在,三次方程的求解公式仍被称为卡尔达诺公式.


HIERONYMI CAR
DANI, PRÆSTANTISSIMI MATHE-
MATICI, PHILOSOPHI, AC MEDICI,
ARTIS MAGNÆ,
SIVE DE REGVLIS ALGEBRAICIS,
Lib. unus. Qui & totius operis de Arithmetica, quod
OPVS PERFECTVM
inscripsit, est in ordine Decimus.

HAbes in hoc libro, studiose Lector, Regulas Algebraicas (Itali, de la Cossa uocant) nouis adinuentionibus, ac demonstrationibus ab Authore ita locupletatas, ut pro pauculis antea uulgò tritis, iam septuaginta euaserint. Neque solum, ubi unus numerus alteri, aut duo uni, uerum etiam, ubi duo duobus, aut tres uni æquales fuerint, nodum explicant. Hunc aût librum ideo seorsim edere placuit, ut hoc abstrusissimo, & planè inexhausto totius Arithmeticæ thesauro in lucem eruto, & quasi in theatro quodam omnibus ad spectandum exposito, Lectores incitarentur, ut reliquos Operis Perfecti libros, qui per Tomos edentur, tanto auidius amplectantur, ac minore fastidio perdiscant.

图 9.4　卡尔达诺《大术》的扉页.(出处:Smithsonian Institution Libraries, Photo No. 76 - 15322.)

下面按照塔尔塔利亚的诗文和卡尔达诺《大术》中的叙述,对三次公式作以详细分析.对于方程,即卡尔达诺所说的“三次和一次幂等于常数”来说:“将一次项系数的 1/3 立方;加上常数项之半的平方;并取整个的平方根.现在对这个数分别加上及减去常数项之半 …,那么,第一个的立方根与第二个的立方根之差便是未知数的值.”[20] 塔尔塔利亚的诗告诉我们,要求两个数 u,v,满足 $u-v=d$(第 3 - 4 行)及 $uv=(c/3)^3$(第 6 - 8 行).因此,$x=\sqrt[3]{u}-\sqrt[3]{v}$(第 10 - 12 行).卡尔达诺作了进一步解释.如果解关于 u,v 的方程组,将 $v=\left(\frac{c}{3}\right)^3\frac{1}{u}$ 代入第一个方程,便得

$$u-\left(\frac{c}{3}\right)^3\frac{1}{u}=d \quad 或 \quad u^2-\left(\frac{c}{3}\right)^3=du,$$

这是一个关于 u 的二次方程,易知

$$u=\sqrt{\left(\frac{d}{2}\right)^2+\left(\frac{c}{3}\right)^3}+\frac{d}{2}.$$

v 的解仅差一个负号.因此,卡尔达诺所阐明的公式为

$$x=\sqrt[3]{\sqrt{\left(\frac{d}{2}\right)^2+\left(\frac{c}{3}\right)^3}+\frac{d}{2}}-\sqrt[3]{\sqrt{\left(\frac{d}{2}\right)^2+\left(\frac{c}{3}\right)^3}-\frac{d}{2}}.$$

卡尔达诺用一种几何推理证明了这个结果.其证明的实质通过下述代数推理能被更容易地看出:

$$\begin{aligned}x^3+cx&=(\sqrt[3]{u}-\sqrt[3]{v})^3+c(\sqrt[3]{u}-\sqrt[3]{v})\\&=u-3\sqrt[3]{u^2v}+3\sqrt[3]{uv^2}-v+3\sqrt[3]{uv}(\sqrt[3]{u}-\sqrt[3]{v})\\&=u-v-3\sqrt[3]{uv}(\sqrt[3]{u}-\sqrt[3]{v})+3\sqrt[3]{uv}(\sqrt[3]{u}-\sqrt[3]{v})\\&=d.\end{aligned}$$

卡尔达诺用例题 $x^3+6x=20$ 来阐明他的法则.因为 $c/3=2$ 且 $d/2=10$,由公式便知

$$x=\sqrt[3]{\sqrt{108}+10}-\sqrt[3]{\sqrt{108}-10}.$$

(在这个公式中,平方和立方根符号及运算符是现代的.卡尔达诺本人是用 R 表示平方根,$cub\ R$ 表示立方根,p 和 m 表示加和减.)关于这里所给出的答案有一个明显的问题.很清楚,方程 $x^3+6x=20$ 的解是 $x=2$.公式所给答案实际上等于 2,但并不那么明显.卡尔达诺曾在几页之后指出这一点,但未具体说明怎样把公式所给答案转化为值 2.

类似地,在"论立方等于一次项和常数",即 $x^3=cx+d$ 的一章中,卡尔达诺呈现和证明了一个稍微不同的法则.在这种情况下,公式变为

$$x=\sqrt[3]{\frac{d}{2}+\sqrt{\left(\frac{d}{2}\right)^2-\left(\frac{c}{3}\right)^3}}+\sqrt[3]{\frac{d}{2}-\sqrt{\left(\frac{d}{2}\right)^2-\left(\frac{c}{3}\right)^3}}.$$

在给出例题 $x^3=6x+40$ 和 $x^3=6x+6$ 之后,他指出,此时若出现 $(c/3)^3>(d/2)^2$ 的困难.那样的话,便不能取其平方根.为了避开这种困难,卡尔达诺对某些特例描述了其它的方法.但是,正像我们随后所知,只是邦贝利才说明了怎样处理卡尔达诺公式中负数的平方根.

卡尔达诺对各种三次方程解法的讨论是在《大术》的第 11—23 章.但是,课本的开始是某些一般性的结果,包括对给定方程的可能根的个数的讨论,而不管根是(真实的)正根,还是(虚构的)负根,还包括怎样用一方程的根确定相关方程的根.例如,卡尔达诺说,这种形式的方程总有一个正解而无负解.相反地,方程的根的个数及符号则依赖于其系数.如果

$$\frac{2c}{3}\sqrt{\frac{c}{3}}=d,$$

那以,这个方程有一个正根 $r=\sqrt{c/3}$ 和一个负根 $-s=-2\sqrt{c/3}$.如果

$$\frac{2c}{3}\sqrt{\frac{c}{3}}>d,$$

那么,有两个正根 r 和 s,及一个负根 $-t$,其中 $t=r+s$.另外,t 是方程 $x^3=cx+d$ 的正根.卡尔达诺还附带指出,在第一种情况下,r 可被看作二重根,因为负根等于 $-2r$.最后,如果

$$\frac{2c}{3}\sqrt{\frac{c}{3}} < d,$$

那么,没有正根.有一个负根 $-s$,其中 s 是 $x^3 = cx + d$ 的正根.沙拉夫丁(Sharaf al - Dīn al - Tūsī)在300多年前对这个方程(及某些其它方程)的根就给出过一个类似的讨论,并对正根的存在性得到过相同的判别标准,但是,是否卡尔达诺也使用了分析极值的相同办法,那就不得而知了.然而,由于他考虑了负根,卡尔达诺确实比他的伊斯兰先辈提供了更多的信息.因此,即使不能证明,卡尔达诺也能知道,什么条件下三次方程有三个实根,它们的和等于 x^2 项的系数.

卡尔达诺的学生费拉里成功地求出四次方程的解.卡尔达诺将这个解简要地叙述在《大术》临结尾部分,其中他列举了20种不同类的四次方程,概述了基本过程,并计算了几个例题.解法开始是一个线性代换以消去 x^3 项,比如,使方程变为了 $x^4 + cx^2 + e = dx$ 的形式.为了解这个方程,给两边加上一个二次项和常数项,使每边成为完全平方式.然后,开平方并计算答案.我们用卡尔达诺的一个例子来解释这个基本过程:$x^4 + 3 = 12x$.如果用 $2bx^2 + b^2 - 3$ 加到两边(其中 b 待定),左边变为 $x^4 + 2bx^2 + b^2$,是一个完全平方式,而右边成为 $2bx^2 + 12x + b^2 - 3$.为使后一式变为完全平方,令 $2b(b^2 - 3) = (12/2)^2$ 或 $2b^3 = 6b + 36$.因此,我们需要解一个关于 b 的三次方程.(这个方程现在叫做所给四次方程的预解方程.)当然,卡尔达诺有解此方程的法则,但此时显然 $b = 3$ 是一个解.因此,所加的多项式是 $6x^2 + 6$,原方程变为了 $x^4 + 6x^2 + 9 = 6x^2 + 12x + 6$.开方,得 $x^2 + 3 = \sqrt{6}(x + 1)$,容易求得解为

$$x = \sqrt{1\frac{1}{2}} \pm \sqrt{\sqrt{6} - 1\frac{1}{2}}.$$

方程的根只有这些吗?也许有人想通过取负平方根求其它根,但这会导致卡尔达诺所忽视的复数值.在其它一些例题中,卡尔达诺确曾用过两组根.也许有人还期望通过预解方程的其它解来求原方程的其它根.卡尔达诺显然分析了这种可能性,但对所发生的情况仅取笑说:“不用我回答是否能找到 b 的另一值 … 来求出[关于 x 的]两个其它解.如果你乐意,自己动手做做看.”[21]

卡尔达诺名著中还有很多有趣之处,包括用负数作为问题的解,以及复数的最早出现.有个问题是,把10分为两部分,使其积为40.用解二次方程的标准技能,卡尔达诺得到两部分为 $5 + \sqrt{-15}$ 和 $5 - \sqrt{-15}$.他将答案检验后确实满足题设条件,但他很困惑,于是写道:“算术就是这样神秘地发展着,这些令人费解的结果真是又精致又不中用.”[22] 因此,卡尔达诺便停止讨论,不再涉及复数.

9.3.2 邦贝利和复数

卡尔达诺的《大术》影响极大,它标志着超越欧洲当时所研究的伊斯兰代数之后首次实质性的进步.作者对自己的著作非常自豪.在课本最后用大写字母写道,“挥毫五载,管用千年.”然而,著作本身很难读懂.推理啰唆,结构上也有许多有待改进之处.为了增进这个课题的传授效果,澄清某些所存在的疑难,在大约15年之后,邦贝利(1526—1572)决定用意大利文写一部更系统的课本来帮助学生学习.虽然在他生前仅出版了五卷中的前三卷,在三次方程多重根的有关问题方面邦贝利还不及卡尔达诺,然而,邦贝利的《代数》仍标志着文艺复兴时期意大利代数的高峰.

邦贝利的《代数》在风格上更接近帕乔利的《集成》和德国人鲁多尔夫的《未知数》,而非卡尔达诺的著作.邦贝利著作的开始是一些基本内容,然后逐渐发展到解三次、四次方程.他与卡尔达诺一样将每类三次方程分开进行论述,但他扩充了卡尔达诺对四次方程的概述,也对四次方程逐类分段解之.在处理完理论材料之后,他利用所建立的方法,给学生呈现了大量例题.他原计划再包含一些

与早期算图著作类似的实用问题，但在研究了梵蒂冈图书馆的丢番图《算术》的一个副本之后，他改变了打算，决定添加一些取材于丢番图及其它著作的抽象数论问题.

人物小传

拉法耶尔·邦贝利(1526—1572)(Rafael Bombelli)

邦贝利早年毕业于工程专业，成人后作为工程师一直服务于其恩主，一位罗马贵族，这位贵族是罗马教皇保罗三世的亲信.他参加的最大项目是将 Val di Chiana 的沼泽地开垦为可耕地.至今，这个位于 Arno 河与 Tiber 河之间宽约 60 英里的山谷仍是意大利中部最肥沃的土地之一.邦贝利后来作为工程顾问参与了罗马附近的 Pontine Marshes 湿地的排水项目.因该地的一场战争在垦荒工作停顿时间，亦即 1557 年到 1560 年的某段时间，邦贝利在其恩主位于罗马的别墅中专心写作他的代数论著.然而，其他的专业事务延迟了著作出版，直到 1572 年他逝世之前才开始排版发行.

代数符号正在替代伊斯兰及早期意大利代数学家的那种严格按照文字的描述，邦贝利对此有所贡献.他用 $R.q.$ 表示平方根，用 $R.c.$ 表示立方根，用一些类似的符号表示更高阶的根.他用括号⌊　⌋包围较长的表达式，如 $R.c.\lfloor 2p\ R.q.21 \rfloor$，但他保留了标准的意大利方式用 p 和 m 表示加和减.他的关键性符号革新是，用半圆上的数字 n 表示未知数的 n 次幂.这样，x^3+6x^2-3x 就被写成 $1\overset{3}{\smile}\ p6\overset{2}{\smile}\ m3\overset{1}{\smile}$.把幂写成数字形式，而非德国的符号形式，使他在单项式乘除法中能容易地表达出指数法则.

在《代数》第一卷的后半部分，邦贝利引入了"另一种和前面非常不同的立方根，它源于论立方等于物共数一章；… 这种根有自己独特的四则运算法则及新的名称."[23] 这种根出现在解三次方程 $x^3=cx+d$ 时，其中$(d/2)^2-(c/3)^3$是负的.邦贝利对这些数拟定了新的名称，它们既不是正的(*più*)，也不是负的(*meno*)，而是现代的虚数.把现代符号所表示的数 bi，$-bi$，邦贝利依次叫做 *più di meno* 和 *meno di meno*.例如，他把 $2+3i$ 写为 $2p\ dim\ 3$，$2-3i$ 写为 $2\ m\ dim\ 3$.邦贝利呈现了关于这些新(复)数的各种乘法法则，诸如，*più di meno* 乘以 *più di meno* 得负，而 *più di meno* 乘以 *meno di meno* 得正$[(bi)(ci)=-bc, bi(-ci)=bc]$.

为了解释他的法则，邦贝利对这些新数的算术四则运算举了大量例子.例如，

$$\sqrt[3]{2+\sqrt{-3}}\cdot\sqrt[3]{2+\sqrt{-3}}=\sqrt[3]{1+\sqrt{-48}},$$

$$\frac{1000}{2+11\,\mathrm{i}}=\frac{1000(2-11\,\mathrm{i})}{(2+11\,\mathrm{i})(2-11\,\mathrm{i})}=8(2-11\,\mathrm{i})=16-88\,\mathrm{i}.$$

加减法法则也详细体现在一些例子中.虽然他认为"整个事情出乎常理，犹如诡辩"，[24] 但毕竟在此首次揭示了复数的运算法则.很明显，邦贝利完全是从已知的实数运算法则中类推和发现这些法则的.类比推理是数学发现的一种基本方法，即使还不能给出严格的证明.由于邦贝利并不知道这些数的真正含义，故不可能给出严格论证.

尽管这样，利用这些复数的法则，针对方程 $x^3=cx+d$，其中$(d/2)^2-(c/3)^3$可以正、也可以是负的，邦贝利还讨论了此时如何使用卡尔达诺公式.他首先分析例题 $x^3=6x+40$.用卡尔达诺公式容易得出 $x=\sqrt[3]{20+\sqrt{392}}+\sqrt[3]{20-\sqrt{392}}$，而答案明显应该是 $x=4$.邦贝利对这两个立方根的和其实就是 4 作了说明.他假设 $20+\sqrt{392}$ 等于某一具有形式 $a+\sqrt{b}$ 的量的立方，即$\sqrt[3]{20+\sqrt{392}}=a+\sqrt{b}$.这意味着$\sqrt[3]{20-\sqrt{392}}=a-\sqrt{b}$.将这两个方程相乘，得 $a^2-b=2$.

进一步,将第一个方程立方并使不含平方根的部分相等,便得 $a^3 + 3ab = 20$.邦贝利并不想用一般的推理求解这个含有两个未知数的方程组,而是指出,由两个方程分别可知 $a^2 > 2$ 且 $a^3 < 20$.满足两个不等式的唯一整数是 $a = 2$.恰好,$b = 2$ 提供了方程组的另一个值,故 $\sqrt[3]{20 + \sqrt{392}} = 2 + \sqrt{2}$.由此便知,三次方程的解可以写为

$$x = (2 + \sqrt{2}) + (2 - \sqrt{2}) = 4.$$

对于例题 $x^3 = 15x + 4$ 来说,由卡尔达诺公式得

$$x = \sqrt[3]{2 + \sqrt{-121}} + \sqrt[3]{2 - \sqrt{-121}},$$

而明显可以看出,答案应该是 $x = 4$.邦贝利使用他新发现的复数知识对此说明如下.首先,假设 $\sqrt[3]{2 + \sqrt{-121}} = a + \sqrt{-b}$,那么 $\sqrt[3]{2 - \sqrt{-121}} = a - \sqrt{-b}$,同上例方法一样,可得两个方程

$$a^2 + b = 5 \quad 和 \quad a^3 - 3ab = 2.$$

因此,邦贝利需要求一个数 a,使 $a^2 < 5$ 且 $a^3 > 2$.他细心地说明 $a = 1$ 和 $a = 3$ 都不成立,因此,$a = 2$ 是唯一的可能.恰好 $b = 1$ 提供了另一个数的解,所以,两个立方根是 $2 \pm \sqrt{-1}$.由此便知,三次方程的解为

$$x = (2 + \sqrt{-1}) + (2 - \sqrt{-1}) = 4.$$

邦贝利针对同一类型的方程还举了一些例子,在每个题中他都能想法算出 a 和 b 的值.然而,他指出,一般的方法是不行的.比如,想用代入法求解关于 a 和 b 的方程组,就会导致另一个三次方程.邦贝利还说明了复数能被用于解以前认为没有解的那些二次方程.例如,他展示了方程 $x^2 + 20 = 8x$ 的解为 $x = 4 \pm 2\,i$.虽然他并不能回答复数使用中的所有问题,但他用复数解决一般问题的实践,给数学家提供了最早的暗示,对复数的研究还是有一定道理和用处的.由于数学家对负数的使用仍然很困惑 —— 卡尔达诺称其为虚构的数,邦贝利根本不认为负数可作为根 —— 那就不应该奇怪为什么需要很多年数学家才真正接受对复数的使用了.

邦贝利是文艺复兴时期的最后一位意大利代数学家.他的《代数》在欧洲其它地区也流传很广.在 17 世纪转折之前,有两位学者,一位在法国,另一位在荷兰,就是通过邦贝利的著作和一些重新发现的古希腊数学著作把代数学推向了新的方向.

9.4　韦达和斯蒂文的工作

16 世纪的欧洲代数学家对于中世纪伊斯兰代数的拓展取得了辉煌成就.尽管他们的符号系统仍有待改进,但其代数操作已很熟练,能够解出直到 4 次的多项式方程.不过,方程的解是以程序规则的形式给出的.未知数及其幂已经符号化,但表示系数的符号尚未出现.因此,对程序的最佳展示就得使用具体的数字例子.这些代数课本中根本没有写出过类似于现今基础代数书中二次求根公式那样的公式.要能写出这样的公式,就得探索新的符号.

16 世纪的意大利除了拓展伊斯兰代数之外,在数学上存在着另一种主要趋势.作为全面复兴古代经典知识的一部分,意大利人对于拯救所有希腊数学著作产生了巨大的兴趣.欧几里得、阿基米德和托勒密的基本著作在几个世纪之前已被翻译.然而,这些翻译者并非专门的数学家,其译文难免有误.对此,16 世纪的学者达成了一致的看法,应该由数学家重新翻译这些著作,还要从原希腊文中翻译出其它的希腊数学著作.在这次数学复兴中最重要的人物是意大利几何学家科曼迪诺(Federigo Commandino,1509—1575),他独自一人用拉丁文翻译了阿基米德、阿波罗尼奥斯、帕普斯、

阿利斯塔克、奥托利科斯、海伦等人的几乎所有著作.科曼迪诺的数学天赋使他得以克服多个世纪以来抄写者们所带入希腊手稿中的种种晦涩.因此,他的每部译作都包含详尽的数学评注,对疑难的澄清以及著作之间的互相引证.

随着(在近古较大图书馆的毁灭中)所幸存下来的希腊数学著作文集,现在整个被呈现于欧洲人面前,他们开始认真地提出这样一个问题,希腊人怎样发现了他们的定理.特别地,由于帕波斯《数学汇编》及其卷Ⅶ,《论分析的领域》的作用,欧洲数学家开始寻找古希腊人所使用的"分析的方法".几乎所有希腊数学著作都是综合推理的模式,从公理出发逐步导向复杂的结果.这些著作,对于其结果是如何得来的,或者怎样才能找到类似的结果,一般并不给出思路分析.帕波斯的著作是唯一一本对希腊几何分析方法提供了一些提示的著作.帕波斯不仅讨论了分析的基本过程,而且在卷Ⅶ中给出了一个希腊主要著作的指南,其中提供了一些工具以便数学家能够用分析法发现新结果或解答新问题.不幸地,除了欧几里得的《已知数》和阿波罗尼奥斯的《圆锥曲线》外,帕波斯所提及的其它著作均已失传.伴随着好奇与挫折,欧洲人就研究所能找到的希腊著作以求搜索出古希腊的方法,并用帕波斯的提示和描述来重构那些失散的著作.

笛卡儿1629年在《思维的指导法则》法则Ⅳ中,极好地表达了这些方面的情感:

> 但是,我后来想到,古代最早的哲学先驱们拒绝承认那些不懂数学的人的学问,他们明显地相信数学是把握其它更重要科学的最简单和必不可少的思维训练和准备.这个时候,我更加确信了我的怀疑,他们曾拥有一种不同于我们这个时代的数学知识.我并不真的以为他们所拥有的这方面知识很完美……但我相信在人类思想中天生就有的某些真理的萌芽…在那个原始而质朴的时代曾有过强大的生命力.……实际上我就是从帕波斯和丢番图的著作中辨识出了这种真正的数学的某些踪迹.……但我的观点是,当时的这些作者以某种不成熟的,令人可叹的方式抑制了这种知识(的传播).很可能他们的做法就像我们所知的许多发现者对待其发现的态度一样;即就是,他们害怕其方法太容易、太简单,以致于泄露出去后会被轻视,他们宁肯取而代之地展示一些空洞无益的真理,在对其艺术的结果进行演绎推理中显示他们足够的天才,以赢得我们对这些成就的敬佩,而不是为我们揭示出那种方法本身,因其会彻底消除我们所赋予的敬佩.终于,在当代涌现出一些天才人物试图复兴这种相同的艺术.由于它正是那门还不正规的"代数"科学,如果我们能把它从无数的数字和令人费解的图形中提炼出来,那么,它将会展示出我们认为真正的数学所应该具有的条理性和简单性.[25]

9.4.1　韦达和《分析的艺术》

最初的"天才人物"之一是韦达(1540—1603),他试图把新代数与希腊的分析方法等同起来,并努力显示这种新代数的"条理性和简单性".他在16世纪最后几年中写了几篇论文,总称为《分析的艺术》,其中通过详细研究方程的结构来替代方程的求解,他卓有成效地重建了代数学研究,从而发展了最早有意识地使用字母的方程理论.

韦达以他1591年的《分析方法入门》(In Artem Analyticem Isagoge)开始了他的计划,并对他所想取得的成就有一个声明:

> 有一种寻求数学真理的方法,据说是由柏拉图最早发现的.赛翁称其为分析法……,尽管古代人仅提出两种分析形式,zetetics和poristics,赛翁对分析的定义也与其完全相符,但我曾加了第三种分析形式,并称之为rhetics或者exegetics.恰当地说,zetetics就是要在

> 某一待定项与若干已知项之间建立方程或比例式；poristics 就是要用方程或比例式检验所述定理的真实性；而 exegetics 则是要在所给方程或比例式中确定出未知项的值. 因此，整个分析艺术在具备这三重功能之后，便可称之为数学中关于正确发现的科学.[26]

回想一下第 5 章的讨论，在那里帕波斯将希腊人的两种分析形式称为问题分析和定理分析. 而在这里韦达不仅改变了名称，还完全改变了其含义，并增加了一种新的分析形式. 对于韦达来说，问题分析变成了 **zetetic 分析**，指把一个问题转化为连接未知量与各已知量的方程的过程；定理分析成了 **poristic 分析**，指通过适当的符号操作揭示定理的真实性的过程；而 **exegetics** 是指对所得的方程进行变形来求出未知数的值的艺术. 韦达将希腊分析等同于代数，并不那么令人奇怪. 在解方程的过程中，人们可以把未知量 x 当作已知量来应用基本运算法则. 经过一系列的运算之后，未知量就能以“$x = \cdots$”的方式用已知量表示出来. 这与帕波斯关于分析的描述具有相同的过程. 但韦达所用术语不同于帕波斯.“求”未知量的实际责任由第三种分析形式(exegetics)来承担，而不是前两种分析形式. 无论如何，韦达的目标是清楚的，这可从前段引文最后一句话以及《分析方法入门》最后一段的文字中看出：“最终，分析的艺术在赋予其三种形式之后，便可宣告一切问题按此框架就都可解了.”[27]

人物小传

弗朗索瓦 · 韦达(1540—1603)(François Viète)

韦达生于法国西部比斯开湾(Biscay)附近的一个村庄，Fontengy-le-Comte. 从普瓦捷(Poitiers)大学法律专业毕业后，他返回家乡开始任律师. 通过和当地名门的联盟，他的法律声望日益增大，并被国王亨利三世召到巴黎任密使，最终于 1580 年在枢密院取得一席位. 宫廷在 1589 年移到都尔(Tours)市之后，他向国王承担的任务之一是作为秘密分析员破解所截获的敌人之间的情报. 他在这方面的研究非常成功，以致被控使用了巫术. 由于他一直在为亨利三世和后来的亨利四世工作，韦达的数学研究只能做为一种业余爱好.

在《分析方法入门》中，韦达的最重要贡献之一是，提出了一种新的符号系统.“算术运算施行于数字；符号运算应用于符号，比如说，字母.”[28] 未知量和已知量都用字母来表示，但为了区别它们，韦达建议“用元音字母表示未知量，而用辅音字母表示已知量.”[29] 虽然他的这个约定不同现今形式，但完全可以由此而对这些字母进行操作了. 而且，这些字母不必仅限制于数字，它们可表示任何满足算术基本运算的量. 然而，韦达并未完全摆脱他的前辈. 他继续使用单词或缩写表示幂，而没有像邦贝利和许凯那样使用指数. 因此，在表示幂的乘法和除法法则时，他就不得不采用文字描述的形式.

在符号运算方面，韦达采用德国的形式用 + 和 - 表示加法和减法. 对于乘法，他通常使用单词 *in*；对于除法，他则用分数线. 因此，

$$\frac{A \text{ in } B}{C \text{ quadratum}}$$

表示现代符号 AB/C^2. 他一般用 L 表示平方根，如 $L64$ 表示 64 的平方根，$LC64$ 则表示 64 的立方根. 像他的大多数前辈一样，韦达也坚持齐性原则，即，在给定的方程中，所有的项必须具有相同的次数. 因此，如果要写的方程是 $x^3 + cx = d$，韦达会认为 c 是面积，而 d 是体积. 他会把此方程写为 *A cubus + C plano in A aequetus D solido*. 注意，他没有用符号，而是用了一个单词，表示“相等”.

尽管韦达刚开始向现代符号系统迈进，但是用字母表示数字常量的决定性步骤，使得他能够脱离其前辈举例的风格和修辞的法则. 他现在已能处理一般的类型而非具体的例子，能够写出公式而

非法则了.另外,符号常量对数字计算的替代,使得人们能把注意力集中到方程的求解程序上,而非具体的解本身.可以看出,求解的程序还适用于数字以外的其它量,比方说,线段或角.进一步,用符号体系求解方程能使解的结构更加明显.比如,在所列公式中保持 $B + D$ 的形式,而不是用 8 代替 $5 + 3$,就可以在求解的最后来分析解与初始常量之间的关系.正是如此,韦达才在某些情况下发现了方程的根与构成该方程的表达式之间的关系.

我们在此分析一下《分析的艺术》一书的各篇论文之中的几个问题及其求解方法,先看《关于符号运算的前言注释》中的材料,这篇论文可能与《分析方法入门》写于同一时期,但直到1631年才出版.在这些注释里,韦达展示了如何对符号量进行操作.他导出了很多标准的代数恒等式,其中大部分至少以文字的形式已为人们所知,但在这里,韦达第一次用纯符号的形式把它们写了出来.比如,韦达注解道,$A - B$ 乘以 $A + B$ 等于 $A^2 - B^2$,$(A + B)^2 - (A - B)^2 = 4AB$.韦达还对从2到6的整数 n 写出了 $(A + B)^n$ 的展开式,但他并没有写出一般的二项式定理.韦达之所以没有揭示出这种一般性的原因,可能与他用单词而非数字表示各种幂有关.类似地,韦达用 $A - B$ 和 $A^2 + AB + B^2, A^3 + A^2B + AB^2 + B^3, \cdots$ 的乘积,分别得到了 $A^3 - B^3, A^4 - B^4, \cdots, A^6 - B^6$,但他未尝试给出一般的法则.

在《前言注释》的另一部分,韦达将他的代数应用于三角学.他用恒等式

$$(BG + DF)^2 + (DG - BF)^2 = (B^2 + D^2)(F^2 + G^2)$$

证明了,给定两个直角三角形,其中一个的底边、垂边和斜边依次为 D、B、Z,另一个的相应边为 G、F、X,则可以构造出一个新直角三角形,其底边是 $DG - BF$,垂边是 $BG + DF$,斜边是 ZX.而且,新三角形底边上的角是原来两个三角形底角的和(假定其和小于90°).由此可知,若从两个全等的直角三角形出发,其底边、垂边和斜边分别为 D、B、Z,则新三角形的三边将相应为 $D^2 - B^2$,$2BD$ 和 Z^2,并且它的底角二倍于原三角形的底角.这些结果等价于三角学中熟知的二倍角公式.接着,韦达用"二倍角三角形"和原三角形进一步构造了"三倍角三角形",其底边为 $D^3 - 3B^2D$,垂边为 $3BD^2—B^3$,斜边为 Z^3.这些关于底边、垂边和斜边的公式等价于现代的三倍角正弦和余弦公式.韦达还用其构造法得出了四倍角和五倍角的公式.

在《Zetetics之五篇》(1591)中,韦达用他的符号算法解答了取自于古代和同时代的各种著作中的大量代数问题.对于每个问题,像所承诺的那样,韦达阐述了怎样导出连接未知量与已知量的方程.他的著作开始是一个与丢番图和约丹努斯的著作开头相同的问题:已知两数的差与和,求这两个数.韦达的程序简明易懂:设差为 B,和为 D,且 A 是两数中较小的数,则 $A + B$ 就是那个较大的数.两数之和就是 $2A + B$,它等于 D.所以,$2A = D - B$,$A = (1/2)D - (1/2)B$.那么另一数 $E = (1/2)D + (1/2)B$.在用符号写出结果之后,韦达又用文字作了重述:"两数和的一半减去差的一半等于较小的数,较小的数加上差就是较大的数".[30] 他以一个例子作为结束:如果 B 是40,D 是100,那么 A 是30,E 是70.这是韦达著作的一种典型格式.尽管他已经引入了符号方法,但他经常用文字重述其结果,好像是想让那些心存疑虑的读者相信,新的符号方法总能转化成更熟悉的文字表达方式.

比较一下丢番图、约丹努斯和韦达处理这个相同问题的不同之处是富有启发作用的.丢番图虽然对于问题作了一般性陈述,但事实上他只解了一个具体的数字例子,即韦达所用的那个例子.约丹努斯一般地解决了这个问题,但用的是文字表达:"从总和中减去两数之差,剩余部分为较小的数的二倍."韦达则完全用的是符号解法.这一问题例证了1350年间代数学之变化.用许多其它的普通问题亦可做出类似的比较.

在Zetetics的第二篇中,韦达讨论了未知值的乘积及其各种幂.他证明了,如果知道两个值的积与平方和,或者两个值的积与和,或者两个值的和与平方差,那么就可以求出这两个值.这一篇中的

几个结果对于韦达后来论述三次方程很重要. 例如, 在问题 20 中, 给定两个值的和与立方和, 要求这两个值. 对这个问题, 韦达令 G 等于未知值的和, D 等于其立方和, 且 A 等于未知值的积. 通过三次二项式的展开式, 韦达导出结果 $G^3 - D = 3GA$, 或者用现代符号表示为

$$(r+s)^3 - (r^3+s^3) = 3(r+s)rs. \tag{9.1}$$

所以, 乘积就可知了, 继而可求出这两个未知值.

9.4.2　韦达的方程理论

韦达在方程理论上的主要贡献出现在《两论方程的整理与修正》一书中. 其中, 韦达呈现了如何把各种类型的方程转化为几个标准型, 以及如何来解每一标准型. 例如, 他不象卡尔达诺和邦贝利那样, 对 13 种三次方程的每一情形分别给出解法过程, 韦达而是先把它们转化为不含二次项的形式, 然后再具体求解. 然而, 在讨论这些公式之前, 韦达阐述了如何用经典的比例论而非根来构造三次方程.

我们具体分析一下方程 $x^3 - mx = n$. 在两论方程之首篇的第 4 章中, 韦达为了保持齐性原则, 将其改写为 $x^3 - b^2x = b^2d$, 并认为这个方程源于四个连比例数, 其中第一个数是 b, 第二个是 x, 而第四个与第二个的差是 d. 即 $b : x = x : y = y : x+d$, 因此

$$\left(\frac{b}{x}\right)^2 = \frac{x}{y}\cdot\frac{y}{x+d} = \frac{x}{x+d}.$$

由此便得 $x^3 = b^2(x+d)$, 即为所给方程. 作为例子, 韦达指出, 如果 $x^3 - 64x = 960$, 那么四个连比例数的第一个是 8, 第四个与第二个的差是 $960/64 = 15$. 因此, 他断言连比例数为 8, 12, 18 和 27, 而方程的解是 $x = 12$.

在第 6 章中, 韦达用另一方式导出了相同的方程. 此处, 他将该方程写为 $x^3 - 3bx = d$, 并用到前面的恒等式(9.1). 如果 r 和 s 的乘积 rs 等于 b, 且 $r^3 + s^3 = d$, 那么 $x = r + s$. 作为例子, 他给出 $x^3 - 6x = 9$, 其中 $6/3 = 2$ 是两数的积, 而两数的立方和是 9. 因此, 这两数是 1 和 2, 由此便知未知数 x 是 3. 韦达进一步指出, 由于 $(r^3 - s^3)^2 = (r^3 + s^3)^2 - 4(rs)^3$, 此法仅适用于 $d^2 > 4b^3$.

如果上述不等式不成立, 有另外两种方法导出该方程. 首先, 韦达用一个更复杂的恒等式, $(r+s)^3 - (r^2+s^2+rs)(r+s) = rs(r+s)$, 来说明如果 $3b = r^2 + s^2 + rs$ 且 $d = rs(r+s)$, 那么 $x = r + s$. 因此, 在例子 $x^3 - 21x = 20$ 中, 他需要求出 r 和 s 使 $r^2 + s^2 + rs = 21$ 且 $rs(r+s) = 20$. 这些值是 1 和 4, 故所求的根为 5.

为了在 $d^2 < 4b^3$ 时对该方程作分析, 韦达用其前面所发现的三角恒等式给出了第二种可选择的方法, 因为"三角学知识可以把这种方程的构造优美而清楚地显示出来."[31] 他将该方程重新写为

$$x^3 - 3b^2x = b^2d. \tag{9.2}$$

那么, 不等式就变为 $b > d/2$. 韦达前面已经指出, 如果原三角形的底边是 D, 垂边是 B, 斜边是 Z, 那么三倍角三角形的底边是 $D^3 - 3B^2D$, 斜边是 Z^3. 用现代术语, 这个底边公式可以重新写为 $\cos 3\alpha = \cos^3\alpha - 3\sin^2\alpha$ 或为 $\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha$, 即

$$\cos^3\alpha - \frac{3}{4}\cos\alpha = \frac{1}{4}\cos 3\alpha. \tag{9.3}$$

令 $x = r\cos\alpha$, 并代入方程(9.2), 可得

$$\cos^3\alpha - \frac{3b^2}{r^2}\cos\alpha = \frac{b^2d}{r^3}. \tag{9.4}$$

比较方程(9.4)和(9.3),首先可得

$$\frac{3b^2}{r^2}=\frac{3}{4} \quad 或 \quad r=2b;$$

其次可得

$$\frac{1}{4}\cos 3\alpha=\frac{b^2d}{r^3}=\frac{b^2d}{8b^3} \quad 或 \quad \cos 3\alpha=\frac{d}{2b}.$$

方程(9. 2)的系数不等式保证最后一式有意义.因此,如果 α 满足 $\cos 3\alpha = d/2b$ 且 $r = 2b$,那么 $x = r\cos\alpha$ 就是方程(9.2)的一个解.例如,如果 $x^3-300x=432$,则 $b=10, d=432/100$.由此可知 $\cos 3\alpha = 432/2000$,查表知 $\cos\alpha = 0.9$,故 $x = 2b\cos\alpha = 18$.

到此为止,如何用代数方法解三次方程的问题仍未解决.对那些相对简单的情况,可以直接使用那些恒等式解之.在两论方程的第二篇中,韦达论述了如何用线性替换 $x = y \pm c$ 将任意的三次方程化简为三种标准型 $x^3+bx=d, x^3-bx=d, bx-x^3=d$ 之一,其中 $3c$ 是原始方程的 x^2 项的系数.因此,利用替换 $x=y+c$ 把方程 $x^3-3cx^2=d$ 就转化为 $y^3-3c^2y=d+2c^3$,而利用替换 $x=y-c$ 把 $x^3+3cx^2-bx=d$ 就转化为 $y^3-(3c^2+b)y=d-2c^3-bc$.在给出这些变换之后,韦达最后对解法作了阐述.比如对于方程 $x^3-3bx=2d$ 来说,韦达首先指出 b^3 必须小于 d^2.(如果这个条件不成立,那么就肯定满足三角方法的条件.)由于 x 可以看成是乘积为 b 的某两数之和,他便建立方程 $y(x-y)=b$,或 $xy-y^2=b$,这个关于 y 的二次方程有两个相异根.从中解出 x,得

$$x=\frac{b+y^2}{y}.$$

代入 $x^3-3bx=2d$,并化简后,得出一个关于 y^3 的二次方程

$$2dy^3-(y^3)^2=b^3.$$

其解为 $y^3=d\pm\sqrt{d^2-b^3}$,从中可看出要求 $b^3<d^2$ 的道理.因为所求的根 x 就是 y 的两值之和,所以,最终的结果是一个和卡尔达诺的描述在形式上稍有不同的公式

$$x=\sqrt[3]{d+\sqrt{d^2-b^3}}+\sqrt[3]{d-\sqrt{d^2-b^3}}.$$

虽然韦达没有考虑方程的负根及复数根,但他在一定程度上论述了根与系数的关系.例如,韦达知道二次方程 $bx-x^2=c$ 有两个正根.为了发现这两个根 x_1 和 x_2 之间的关系,他从 $bx_1-x_1^2=bx_2-x_2^2$ 推知 $x_1+x_2=b$,即"b 是所求的两根之和."用 x_1+x_2 替代方程 $bx_1-x_1^2=c$ 中的 b,他得到 $x_1x_2=c$,即"c 是所求的两根之积."[32]

韦达想对三次方程 $bx-x^3=d$ 作同样的讨论.他也知道这个方程有两个正根 x_1, x_2.在此情况下,结果并不那么简单.韦达发现,系数 b 等于 $x_1^2+x_2^2+x_1x_2$,而常数项 d 是 $x_1^2x_2+x_2^2x_1$.具有三个正根的三次方程的情况会如何?当然,这样的方程肯定含有二次项.韦达的标准解法可以通过线性替换将其化为不含二次项的方程,而新方程最多只能有两个正根.因此,韦达就不能正常地找到第三个根.事实上,在一个例子中,对于方程 $x^3-18x^2+88x=80$,他仅算出化简后的方程 $20y-y^3=16$ 的一个整数根,因此,也只求出原方程的一个根.然而,在两论方程一书的结尾部分,韦达没有证明地陈述了四个命题,每个命题对应着二次到五次中的一个方程,方程的系数都是根的基本对称函数.例如,对于三次方程来说,"如果 $x^3-x^2(b+d+g)+x(bd+bg+dg)=bdg$,则 b,d 和 g 都是方程的根."对于四次方程,"如果 $x(bdg+bdh+bgh+dgh)-x^2(bd+bg+bh+dg+dh+gh)+x^3(b+d+g+h)-x^4=bdgh$,则 b,d,g 和 h 都是方程的根."[33]韦达认为这些定理"非常优美",是该书的"花冠",我们也就在此结束对这位卓越的法国代数学家的研究吧.

9.4.3 斯蒂文和十进制小数

与韦达同时代、但主要生活在荷兰的另一位数学家是斯蒂文(1548—1620),他在 17 世纪转折时期对一种主要的数学思想变化做出了实质性贡献.他经过慎重考虑后创造了十进制小数的符号,并大力提倡对它的应用.他也在改变"数"的基本概念以及消除亚里士多德对数与量的区别方面扮演了关键角色.这些贡献反映在他 1585 年出版的两本重要的数学著作之中.其中一本是《论十进制》(De Thiende),法文版书名叫 La Disme;另一本是《算术》(L'Arithmétique),论述了算术和代数的一般理论.[34]

在中世纪后期及文艺复兴时期,欧洲没有使用十进制分数.虽然在整个 13 到 16 世纪欧洲的各种算术课本中都要讨论印度 — 阿拉伯位值系统,但它仅被用于整数.如果需要分数,则将其写成普通分数的形式,或在许多三角学著作中为六十进制分数.鲁多尔夫和韦达在他们 16 世纪的著作中都暗含了十进制分数,但未有较大影响.例如,韦达指出,要想求出精度较高的 2 的平方根值,就得添加足够多的零并算其平方根,比如,20 000 000 000 000 000 000 000 000 000 000.他算出这个数的根是 141 421 356 237 309 505,因此,2 的平方根近似于 $1\frac{41\ 421\ 356\ 237\ 309\ 505}{100\ 000\ 000\ 000\ 000\ 000}$.鲁多尔夫曾用一条垂直线分开整数部分和小数部分来表示十进分数.比如他用 413|4375 表示 413.4375.但是,不论鲁多尔夫还是韦达,或者也用 10 的幂作分母表示分数的其他人,他们都没有显示出对这种分数概念的深刻理解.然而,在伊斯兰世界中,塞毛艾勒已理解了这个概念,卡西(al-kashi)还创设了十进制分数的一种方便的记号.

人物小传	西蒙·斯蒂文(1548—1620)(Simon Stevin)
	斯蒂文(图 9.5)生于现今比利时的布鲁日(Bruges)的一户富裕家庭,但他后来离开当时处在西班牙统治下的出生地到达荷兰.斯蒂文的大半生服务于荷兰行政长官摩利士将军(Maurice of Nassau).先是作为摩利士的工程师,数学和弹道学教师,以及经济和航海顾问,后从 1593 年起任荷兰政府的军需局长官.应摩利士请求,他在莱顿(Leiden)大学成立一个工程学院,并用荷兰语而非拉丁语教学.为了满足国家对工程师,商人,勘测员,和航海员的急需,斯蒂文用荷兰语对于相关课程写了大量教材.



图 9.5　比利时邮票上的斯蒂文.

斯蒂文并未受这种伊斯兰式发展的影响,他在《论十进制》中将其想法和符号有机地溶于一体.在前言中斯蒂文陈述了其著作的目的:"简而言之,本书要讲授不使用普通分数怎样简便地完成各种帐务结算,数字计算和货币兑换;如果这种新记数系统的算术运算法则与整数的相关法则相通,便能起到这些效果."[35] 这就是说,斯蒂文保证在他的新系统中的所有运算能完全像整数中那样被完成.当然,这是十进制小数的基本优点.在《论十进制》的开头,斯蒂文用印度——阿拉伯数字,将"十进制"定义成以 10 为倍数的几何级数上的算术,并称整数部分为"起点",记为⓪.因此,364 被写为 364⓪.他的主要定义是第三个,其中描述了小数部分的术语和符号:"我们把起点单位的十分之一称之为首位,记为 ①,把首位单位的十分之一称之为第二位,记为 ②,依次类推,总是在前一位号的基础上再加上一."[36] 他举例解释道:3①7②5③9④ 意味着 3/10,7/100,5/1000,9/10 000,或者总体为 3759/10 000.类似地,$8\frac{937}{1000}$ 被写为 8⓪9①3②7③.斯蒂文指出,在他的记数系统中不出现分

数形式;除了位号 ⓪ 外,每个位号(带圆圈的数字)左边只有一位数字;按照这些法则写出的数叫做**十进制数**.

在小册子的第二部分,斯蒂文讲述了十进制数的基本运算.实际上,其重要思想是这种数系的运算完全和整数运算相同,但须考虑适当的位号.因此,在加法和减法中所有数的 ① 号位必须上下对齐.对于乘法,斯蒂文指出,先按整数相乘,其乘积最右边数字的位号是被乘数与乘数最右边数字的位号之和.对于除法,只须类似地从被除数最右边的位号中减去除数最右边的位号.他还给出了求平方和立方根时确定位号的法则.因此,虽然他的符号有点不同于现代形式,但斯蒂文已经建立了用十进制小数进行运算的基本法则和理论基础.在《论十进制》的最后总结部分,他提倡在各种贸易计算中使用他的新十进制系统.他建议在每种情况下都用一个已知的基本单位作为起点位,然后用他的系统表示该单位的分数部分.然而,他的建议直到 200 年之后,当法国的革命政府引入公制时,才被广泛实行.

《论十进制》中的十进制小数与数的基本概念的变化相关联,这一思想在斯蒂文 1585 年的另一部数学著作《算术》中得以充分体现.当然,在他之前的数个世纪中,已有许多数学家把无理量看作"数",亦即,用与整数一样的法则和概念在处理无理量.逐渐地,区分数与量的欧几里得传统便被打破.而斯蒂文是最早明确宣告这一点的人,在《算术》开头有两个定义:

1. 算术是数的科学.
2. 数可用于解释一切事物的量.

这说明在著作的开始,斯蒂文就指出数能表示量,任何类型的一切量.数不再像欧几里得定义的那样仅是单位的集成.斯蒂文甚至在那一页的上方用大写字母写道,那个单位元素本身也是一个数.古希腊人是否认这一思想的.对于他们,单位元素本身不是数,仅是数的生成元,正像点是线的生成元一样.多少世纪以来,这种思想一直受到争论.晚到 1547 年,费拉里在前面提到的公开挑战中送给塔尔塔利亚的问题之一便是,单位元素是否一个数.塔尔塔利亚抱怨这个问题不关乎数学,而是玄学.然后他模棱两可地断言,单位元素是一个"潜在的"数,而非"实际上的"数.相对而言,斯蒂文则非常肯定.他哲学上的基本推理是,部分属于全体,而单位元素是若干单位(即,"数")的一部分,故其本身是数.数学推理只不过是,可以对单位元素像对其它"数"一样地进行运算.特别地,可把单位元素分成任意多的部分.欧几里得将单位元素专门作为"单位集成"的基础,故而作为区分离散与连续的基础,不再对斯蒂文有意义.他还大胆断言"数不是非连续的量."[37] 任何量,包括单位,可被"连续地"划分.从某种意义上讲,这正是十进制小数的思想基础.可以对单位元素进行任意细的划分来连续得到一系列尽可能多的位号.

斯蒂文给出几个专门的定义进一步解释了数的含义.例如,"解释几何量值的数称之为几何数,并获得相应名称进一步确定所解释量的种类.""平方数"表示平方,"立方数"表示立方.然而,斯蒂文指出,任何(正)数都是平方数,故其平方根也是数:"部分属于全体,8 的根是其平方 8 的部分.故 $\sqrt{8}$ 属于 8 的同类事物.而 8 是数,故 $\sqrt{8}$ 也是数."[38] 然后,《论十进制》的十进制数系统使斯蒂文能够像表示 8 本身一样,把 $\sqrt{8}$ 表示到任意所需的精度.

斯蒂文对可公度的数和不可公度的数作了辨识,但在他思想中所有这些量都是数.因此,欧几里得在卷 X 中对各种无理线段的区分就失去了意义.所有这些线段都可表示为数并能按标准的算术运算处理.斯蒂文指出,通过取更多的根及根的组合,我们甚至可比欧几里得得到更多种类的线段;而所有这些线段(或数)都可用他的十进算术算之.

站在现在的高度,离散的欧几里得"数"早已被混合成连续的数轴,故有点难以理解斯蒂文的

基础性贡献.但是,欧几里得一直是数学研究的中心,他的思想大家都得正视.如果想改变他的理论,就需要长久不懈的斗争.长久以来,欧洲和伊斯兰世界阅读欧几里得的许多数学家,他们都没有顾及这种区别.特别是本章所研究的数学家,他们试图用同一方式处理所有的量.然而,那些具有哲学倾向的学者因这种对欧几里得著作的傲慢态度而烦恼.数学家需要进一步证实这种区别在数学上不再必要.自然地,斯蒂文单独一人不可能做到.直到19世纪,将“离散的算术”置人“连续的量”这一工作才告完成.然而,斯蒂文处于数学思想的转折点.他是如此成功,人们最终已难以理解以前所发生的一切.

习　题

意大利算图课本中的问题

1.在意大利卢卡,1金币值5里拉,12索尔多,6迪拉里尼斯.问13索尔多,9迪拉里尼斯值多少金币?(注意,20索尔多为1里拉,12迪拉里尼斯为1索尔多.)

2.若8braccia长的布料值11金币,则97braccia长的布料值多少?

3.今有合成银25磅其中每磅含纯银8盎司,另有合成银16磅每磅含银$9\frac{1}{2}$盎司.问合并后还需加入多少铜,才能铸成每磅含银$7\frac{1}{2}$盎司的硬币?

4.印刷最早的算术课本Treviso Arithmetic(1478)中的问题:罗马主教派一信使到威尼斯,令其7天到达,威尼斯教会也派一信使到罗马,令其9天到达,两地相距250哩.若两个信使恰巧同时出发,问他俩几天相遇,各走了多少哩?

5.本题及6、7题源于弗兰切斯卡的著作.今有三人合股,第一人投资58金币,第二人87金币,但不知道第三人的投资数.他们的总收入是368,其中第一人分得86.问另两人应该各得多少,第三人投资多少?

6.今有三人完成一项任务,第二和第三人需要10天,第一和第三人需要12天,第一和第二人需要15天.问三人单独完成这项任务,各需要几天?

7.某一喷水池有上下两个盆,每个盆各有三个出口.上盆的第一出口注满下盆需要2小时,第二出口需要3小时,第三出口需要4小时.当上面三个出口全部关闭时,下盆的第一出口抽空它需要3小时,第二出口需要4小时,第三出口需要5小时.如果所有出口都打开,问充满下盆需要多久?

8.安托里奥著作的问题:求两个数,使其乘积为8且平方和为27.(令第一个数为$x+\sqrt{y}$且第二个数为$x-\sqrt{y}$,则有$x^2-y=8$,且$2x^2+2y=27$.)

9.把10分为两部分,使97减去第一部分的平方,再开方;100减去第二部分的平方,再开方;上述两平方根的和为17.(此题也源于安托里奥的著作.安托里奥设两部分为$5+x$和$5-x$并导出一个关于x的方程.)

10.达尔迪指出某些四次方程$x^4+bx^3+cx^2+\mathrm{d}x=\mathrm{e}$的解为

$$x=\sqrt[4]{\left(\frac{d}{b}\right)^2+e}-\sqrt{\frac{d}{b}}.$$

他描述解法的例子如下:某人租借给别人100里拉,4年后的本利和为160里拉,其利息每年按复利计.利率是多少?象课本中的例题一样,设1里拉每月的利率是X迪拉里尼斯.证明这个问题可以导出方程$x^4+80x^3+2400x^2+32\,000x=96\,000$,且上述公式所给的解就是通过完全四次方程和公式求出的解.

11.P·弗兰切斯卡给出这样的问题:把10分为两部分,使它们的积除以它们的差,其结果为$\sqrt{18}$.在解题时,他用到四次方程$ax+bx^2+cx^4=d+ex^3$的解:

$$x=\sqrt[4]{\left(\frac{b}{4c}\right)^2+\frac{d}{c}}+\frac{e}{4c}-\sqrt{\frac{a}{2e}}.$$

试说明这个公式在本题中有效,但一般不成立.弗兰切斯卡是如何导出这个公式的?

12. L·帕乔利在《算术、几何、比与比例集成》中关于方程 $6x^3 = 43x^2 + 79x + 30$ 的解法如下："该常数与诸物相加得另一常数，在化简为1立方幂之后(即所有项除以6)，那么，1立方幂等于 $7\frac{1}{6}$ 平方幂加 $18\frac{1}{6}$. 然后，将诸平方幂分为一半，把这一半自乘，并加 $18\frac{1}{6}$. 则结果为 $31\frac{1}{144}$，且1物便等于这个结果的根加 $3\frac{7}{12}$，其中 $3\frac{7}{12}$ 是诸平方幂之半." 试说明帕乔利的答案不正确. 在展现其解法时，他在联想些什么？

许凯的问题

13. 用许凯的逼近算法求 $\sqrt{6}$ 的近似值. 亦即，因为 $2\frac{4}{9} < \sqrt{6} < 2\frac{5}{11}$，所以下一步取近似值 $2\frac{9}{20}$. 继续这个过程，直到得出许凯的最佳近似值 $2\frac{89}{198}$.
14. 用许凯的逼近算法，求出他所得 $\sqrt{5}$ 的近似值 $2\frac{161}{198}$.
15. 两数之比为5：7，且较小数的平方乘以较大数为40. 求这两数.
16. 今有一数乘以20再加7，其和与另一数之比为3：10，另一数是原数乘以30再减9. 求这个数.（许凯认为此题无解. 为什么？）
17. 在某一盛满酒的容器上有三个龙头，大龙头3小时抽空该容器，中龙头4小时，小龙头6小时. 如果三个龙头都打开，问多久抽空该容器？
18. 某丈夫死时妻子身怀有孕，他遗赠100法郎如下：如果生下女儿，母亲得女儿的两倍. 如果生下儿子，儿子得母亲的两倍.（性别歧视问题！）但后来母亲生下一儿一女双胞胎. 根据父亲的意愿，该如何处置遗产？

鲁多尔夫《未知数》的问题

19. 将 $\sqrt{27 + \sqrt{200}}$ 表示为 $a + \sqrt{b}$ 的形式.
20. 某人欠款3240金币，第1天支付1金币，第2天2金币，第3天3金币，等等. 问还清欠款需要多少天？
21. 把10分为两部分，使其乘积等于 $13 + \sqrt{128}$.

施蒂费尔《整数算术》的问题

22. 在奇数数列中，第一个奇数等于 1^5. 跳过1个数后，随后4个数之和 $(5 + 7 + 9 + 11)$ 等于 2^5. 再跳过后3个数后，随后9个数之和 $(19 + 21 + 23 + 25 + 27 + 29 + 31 + 33 + 35)$ 等于 3^5. 在每一相继步骤，跳过后继的三角形数个奇数. 试用现代记号将这个5次幂的幂规律公式化，且用归纳法证之.
23. 斯蒂费尔（以及朔伊贝尔和其他的同时代人）求高阶根的方法基础是相应的二项展开式，或更明确地，是帕斯卡三角形的相关行的表值. 例如，欲求1 336 336的四次方根，首先应该说明答案是一个3开头的两位数. 然后从原数中减去 $30^4 = 810\,000$，得余数526 336. 考虑到帕斯卡三角形第四行的表值为1，4，6，4，1，猜想下一位数字为4，并从该余数中相继减去 $4 \times 30^3 \times 4 = 432\,000$，$6 \times 30^2 \times 4^2 = 86\,400$，$4 \times 30 \times 4^3 = 7680$，和 $4^4 = 256$. 这样，其结果为0，故所求根为34. 请书面详细解释这个过程，且用它计算10 556 001的四次方根.

雷科德《智力磨石》的问题

24. 某个军队由公爵，伯爵和士兵组成. 每位公爵手下的伯爵人数的二倍等于公爵人数. 每位伯爵手下的士兵人数的四倍等于公爵人数. 士兵人数的二百分之一等于公爵人数的九倍. 问各有多少人？
25. 某绅士考问一位好自夸的算学家，说：我两手共有8钱，每个手中的钱数加上它的平方数和立方数，总和为194. 问每个手各有几钱？

卡尔达诺《大术》的问题

26. 证明:如果 r,s 是 $x^3+d=cx$ 的两个正根,那么 $t=r+s$ 是 $x^3=cx+d$ 的一个根.

27. 证明:如果 t 是 $x^3=cx+d$ 的一个根,那么

$$r=\frac{t}{2}+\sqrt{c-3\left(\frac{t}{2}\right)^2} \quad 和 \quad S=\frac{t}{2}-\sqrt{c-3\left(\frac{t}{2}\right)^2}$$

都是 $x^3+d=cx$ 的根.应用这个法则解方程 $x^3+3=8x$.

28. 证明方程 $x^3+cx=d$ 总有一个正解而无负解.

29. 利用卡尔达诺的公式解方程 $x^3+3x=10$.

30. 利用卡尔达诺的公式解方程 $x^3=6x+6$.

31. 分析方程 $x^3=cx+d$.证明当$(c/3)^3>(d/2)^2$时,方程有三个正解.(而此时卡尔达诺的公式会出现虚量.)

32. 解方程 $x^3+21x=9x^2+5$,首先用替换 $x=y+3$ 消去 x^2 项,然后解出关于 y 的导出方程.

33. 用费拉里的方法解四次方程 $x^4+4x+8=10x^2$.先变形为 $x^4=10x^2-4x-8$,两边同时加上 $-2bx^2+b^2$.确定 b 应满足的三次方程,以便导出的方程两边为完全平方形式.对于三次方程的每一个解,求出 x 的所有解.题设方程共有几个不同的解?

34. 弗朗西斯妻子的嫁妆比弗郎西斯自己的财产多100奥里斯,且嫁妆钱数的平方比他的财产钱数的平方多400.求该嫁妆和财产.(注意财产的答案是负的.卡尔达诺解释为债务.)

费拉里和塔尔塔利亚竞赛中的挑战题

35. 求两个数 x 和 y,其中 $x>y$,使 $x+y=y^3+3yx^2$ 且 $x^3+3xy^2=x+y+64$,塔尔塔利亚的解是 $y=x-4$,而

$$x=\sqrt[3]{4+\sqrt{15\frac{215}{216}}}+\sqrt[3]{4-\sqrt{15\frac{215}{216}}}+2.$$

36. 把8分成 x 和 y 两部分,使 $xy(x-y)$ 取最大值.

邦贝利的问题

37. 方程 $x^3+3x=36$ 的一个根显然是3,而用卡尔达诺公式算出的根是:

$$x=\sqrt[3]{\sqrt{325}+18}-\sqrt[3]{\sqrt{325}-18}$$

试用邦贝利的方法,证明这个数事实上等于3.

38. 把$\sqrt[3]{52+\sqrt{-2209}}$ 表示为 $a+b\sqrt{-1}$ 的形式.

韦达的问题

39. 试证明韦达的下述定理:如果给定两个直角三角形,其中一个的底边、垂边和斜边依次为 D、B、Z,另一个的相应边为 G、F、X,则可以构造出一个新直角三角形,其底边是 $DG-BF$,垂边是 $BG+DF$,斜边是 ZX.而且,新三角形的底角等于原来两个三角形底角的和.

40. 给定两数的乘积及其比率,求这两数.设这两数是 A,E,且 $AE=B$,$A:E=S:R$.证明 $R:S=B:A^2$,且 $S:R=B:E^2$.韦达的例题之一是 $B=20,R=1,S=5$,试说明 $A=10$ 且 $E=2$.

41. 给定两数的差及其立方差,求这两数.设 E 是这两数的和,B 是它们的差,D 是其立方差,证明 $E^2=(4D-B^3)/3B$.知道 E^2 和 E 之后,这两数本身便可知.当 $B=6$ 和 $D=504$ 时,求其解.

42. 试说明方程 $x^3+b^2x=b^2c$ 源于四个连比例数,其中第一个数是 b,第二个数是 x,而第二个与第四个的和是 c.利用这个结果解方程 $x^3+64x=2496$.

斯蒂文的问题

43.将 13.395 和 22.864 2 写成斯蒂文的记数形式.利用他的法则将这两数相乘并用第一个去除第二个.

44.给定两数 237⓪5①7②8③ 和 59⓪7①3②9③,求其差.

专题讨论

45.为什么在大学代数课中,一般不再讲授卡尔达诺公式?应该讲授它吗?它能给方程理论研究带来什么见解?

46.通过那些使用卡尔达诺公式时出现的具体问题,即把一个实数根给定成了两个复值的和,来拟定一个引进复数研究的教学提纲.讨论这种处理方式的优点.

47.比较书中所讨论的,数学家曾用过的未知数的各种符号形式.写出一篇短文,论先进的符号表示在提高学生的代数理解能力上的作用.

48.最早印刷出版的算术书是一位无名作者的 Treviso Arithmetic(1478).写一篇短文,论该书的内容及重要性.查阅注释 39 中的书:Frank J. Swetz, Capitalism and Arithmetic.

49.为什么数学知识对于文艺复兴时期的商人是必要的?他们确实需要了解三次方程的解吗?那么,在 16 世纪后期的课本中,详细研究这些方程的目的是什么?

50.比较约丹努斯和韦达的符号体系.韦达的著作在哪些方面比约丹努斯的更先进?

51.解释为什么 16 世纪的数学家,将新代数等同于帕波斯所描述的希腊分析.

文献和注解

关于这一章内容的综合性著作有:Paul Lawrence Rose, The Italian Renaissance of Mathematics: Studies on Humanists and Mathematicians from Petrarch to Galileo(Geneva: Droz, 1975), Warren van Egmond: "The Commercial Revolution and the Beginnings of Western Mathematics in Renaissance Florence, 1300—1500," Dissertation(University of Indiana, 1976) and R. Franci and L. Toti Rigatelli, "Towards a History of Algebra from Leonardo of Pisa to Luca Pacioli," Janus 72(1985), 17 – 82. Chapter 2 of B. L. Van der Waerden, A History of Algebra from al – Khwarizmi to Emmy Noether (New York: Springer, 1985).

1.引自 R. Franci and L. Toti Rigatelli, "Towards a History of Algebra," pp. 64 – 65.

2. Girolamo Cardano, *The Great Art, or The Rules of Algebra*, translated and edited by T. Richard Witmer(Cambridge: MIT Press, 1968), p. 96.

3. Warren van Egmond, "Commercial Revolution." 这篇学位论文检查了各种 *maestri d'abbaco* 著作,含算术和代数.它也论述了这些著作将基本的数学思想重新引入综合文化的极端重要性.

4. R. Franci and L. Toti Rigatelli, "Towards a History of Algebra," p. 31. 这篇文章分析了各种 *maestri d'abbaco* 著作与斐波那契和帕乔利著作的关系.

5. R. Franci and L. Toti Rigatelli, "Fourteenth – century Italian algebra," in Cynthia Hay, ed., *Mathematics from Manuscript to Print*: 1300—1600(Oxford: Clarendon Press. 1988), 11 – 29, p. 16. 这篇文章概述了 14 世纪的几本重要算图手稿的内容.

6. R. Franci and L. Toti Rigatelli, "Towards a History of Algebra," p. 49.

7. Warren van Egmond, "Commercial Revolution," p. 266. 有关 Mazzinghi 的内容参阅:R. Franci and L. Toti Rigatelli, "Towards a History of Algebra," and B. L. van der Waerden, *History of Algebra*.

8. Warren Van Egmond, "The Algebra of Master Dardi of Pisa," *Historia Mathematica* 10(1983), 399 – 421.

9. Graham Flegg, Cynthia Hay, and Barbara Moss, *Nicolas Chuquet, Renaissance Mathematician*(Boston: Reidel, 1985) 中翻译

了许凯的手稿,并讨论了作者的生平及他在数学史中的地位.

10. 同上,p.90.

11. 同上,p.105.

12. 同上,p.151.

13. 同上,p.177.

14. 鲁多尔夫 1525 年的《未知数》没有现代版本.关于德国文艺复兴时期的所有代数著作,参见:P. Treutlein, "Die Deutsche Coss," *Abhandlungen zur Geschichte der mathematischen Wissenschaften* 2(1879), 1 – 124. A more recent work, a collection of articles by experts on individual German algebraists and Rechenmeisters, is Rainer Gebhardt and Helmuth Albrecht, eds., *Rechenmeister und Cossisten der frühen Neuzeit* (Annaberg-Buchholz: Schriften des Adam-Ries-Bundes, 1996).

15. Joseph Hofmann, *Michael Stifel* 1487? – 1567: *Leben, Wirken und Bedeutung für die Mathematik seiner Zeit* (Wiesbaden: Franz Steiner Verlag, 1968).中讨论了施蒂费尔的《整数算术》.

16. 影印本见 Robert Recorde, *The Whetstone of Witte* (New York: Da Capo Press, 1969). 没页码.

17. 有关努涅斯的著作介绍,生平,及他的代数著作影印本,分别参见:H. Bosmans, "Sur le 'Libro de algebra' de Pedro Nuñez" *Bibliotheca Mathematica* (3)8(1907), 154 – 169 and H. Bosmans, "L'Algebre de Pedro Nuñez," *Annaes scientificos da academia politechnica do Porto* 3(1908), 222 – 271. A more general treatment of his life and work is by Rodolpho Guimarñes. *Sur la vie et l'oeuvre de Pedro Nuñes* (Coimbra: Imprimerie l'Université, 1915). The text of the Libro de Algebra was photographically reprinted by the Academia das Ciências de Lisboa in 1946.

18. John Fauvel and Jeremy Gray, ed., *The History of Mathematics: A Reader* (London: Macmillan, 1987), p.254.这本书的第 8A 章翻译了费罗和塔尔塔利亚以及费拉里和塔尔塔利亚的挑战中的大多数问题,还翻译了塔尔塔利亚对他和卡尔达诺之间的讨论的描述.

19. 由我女儿 Sharon Katz 译自意大利文.

20. Cardano, *The Great Art*, p.98 – 99.这本书应认真阅读,其中有许多珍品本书未讨论.有关卡尔达诺的英文传记及他本人的自传的英译本,分别参见:Oystein Ore, *Cardano: The Gambling Scholar* (Princeton: Princeton University Press, 1953). *Cardano, The Book of My Life*, translated by J. Stoner (New York: Dover, 1962). Richard Feldman, "The Cardano-Tartaglia Dispute," Mathematics Teacher 54(1961), 160 – 163 and James Bidwell and Bernard Lange, "Girolamo Cardano: A Defense of His Character," Mathematics Teacher 64(1971), 25 – 31.

21. 同上,p.250.

22. 同上,p.220.

23. Rafael Bombelli, Algebra (Milan: Feltrinelli, 1996), p.133.这是 1572 年原书的现代意大利文版,但没有英语译文.有关邦贝利的详细情况,参见如下三篇文章:S. A. Jayawardene: "The Influence of Practical Arithmetics on the Algebra of Rafael Bombelli," Isis 64(1973), 510 – 523; "Unpublished Documents Relating to Rafael Bombelli in the Archives of Bologna" Isis 54(1963), 391 – 395; and "Rafael Bombelli, Engineer – Architect: Some Unpublished Documents of the Apostolic Camera," Isis 56(1965), 298 – 306.

24. 同上.

25. René Descartés, *Rules for the Direction of the Mind*, translated byElizabeth S. Haldane and G. R. T. Ross, Great Books edition (Chicago: Encyclopedia Britannica, 1952), pp.6 – 7.

26. François Viète, The Analytic Art, translated and edited by T. Ricchard Witmer (Kent, Ohio: Kent State University Press, 1983), pp.11 – 12. 后面所有引文都取自这个版本.关于《两论方程的整理与修理》,另有一个更忠实原著的版本(当然,现代读者就更难读懂)如下,其中含有许多本书未讨论的韦达著作:by Robert Schmidt, published in *The Early Theory of Equations: On Their Nature and Constitution* (Annapolis: Golden Hind Press, 1986).

27. 同上,p.32.

28. 同上,p.17.

29. 同上,p.24.

30. 同上,p.84.
31. 同上,p.174.
32. 同上,p.210.
33. 同上,p.310.
34. Charles Jones:"On the Concept of One as a Number"(University of Toronto,1978)的第二部分主要研究斯蒂文十进制小数的思想及对希腊人数的概念的打破,本书引用了其中的观点.
35. Henrietta O. Midonick. ed., *The Treasury of Mathematics*(New Youk: Philosophical Library,1965),p.737.有关斯蒂文的详细情况,见:Dirk J. Struik, The Land of Stevin and Huygens: *A Sketch of Science and Technology in the Dutch Republic During the Golden Century*(Dordrecht: Reidel,1981).
36. 同上,p.740.
37. Charles Jones,"Concept of One,"p.239.
38. 同上,p.248.
39. Frand J. Swetz, *Capitalism and Arithmetic: The New Math of the* 15 *Century*(La Salle, Ⅱ: Open Court,1987),p.158.这本书完整翻译了 Treviso Arithmetic,并详细阐述了原著产生的数学和社会背景.
40. R. Franci and L. Toti Rigatelli,"Towards a History of Algebra,"p.65.
41. Robert Recorde, *The Whetstone of Witte*.

文艺复兴时期代数学概览

14 世纪中叶	比萨的达笛(Maestro Dardi of Pisa)	算盘书
1353—1383	德·马志尼(Antonio de' Mazzinghi)	算盘书
约 1431—1487	许凯(Nicolas Chuquet)	法国代数学
1445—1517	帕乔利(Luca Pacioli)	意大利算术和代数学
1465—1526	费罗(Scipione del Ferro)	三次方程
16 世纪初	鲁道夫(Christoff Rudolff)	德国 *Coss*
1487—1567	斯蒂弗尔(Michael Stifel)	德国 *Coss*
1494—1570	舒贝尔(Johann Scheubel)	根的计算
1499—1557	塔塔利亚(Niccolò Tartaglia)	三次方程
1500—1558	内夫(Annibale della Nave)	三次方程
1501—1576	卡尔丹(Gerolamo Cardano)	三次方程
1502—1572	鲁尼斯(Pedro Nunes)	葡萄牙代数学
1509—1575	柯门第诺(Federigo Commandino)	希腊著作的翻译
1510—1558	雷科德(Robert Recorde)	英国代数学
1522—1565	费拉里(Lodovico Ferrari)	四次方程
1526—1572	邦贝里(Rafael Bombelli)	复数
1540—1603	韦达(François Viète)	方程理论
1548—1620	斯蒂文(Simon Stevin)	十进制分数

第10章 文艺复兴时期的数学方法

> 没有什么(上帝的作品除外)会像真正的艺术与科学那样使人们的心灵变得美丽多彩.……在诸多……可以美化人们心灵的学问中,数学可以说独占鳌头.然而,如果不了解几何原理,不学习几何《原本》,任何人都不可能掌握数学的知识.
>
> ——约翰.迪伊:《数学序言》[1]

1616年2月26日,罗马教皇的顾问、宗教裁判所红衣主教贝拉迈(Bellarmine)派了两名官员将伽利略遣送回乡.三个月后,这位红衣主教给伽利略寄来一纸可以反映当时情况的文书:"伽利略先生台鉴,……教皇[保罗五世]陛下假《禁书全目》公谕:日心说与圣经相悖,不得信奉或为之辩护."[2]

代数并不是文艺复兴时期数学家们仅有的关注对象.事实上,正如约翰·迪伊在本章章头引言中所说的那样,几何学仍然是数学的中心课题.作为对古典学术兴趣普遍复活的一部分,文艺复兴时期的学者们也开展了对希腊几何的研究.首先是对欧几里得的研究,他的有多种拉丁文译本的《原本》,是当时欧洲大学里的主要数学课程.据说当时任何以博学自居的人对欧几里得的著作都会有所了解.

因为有许多人不懂拉丁文或没有上过大学,在16世纪便出现了《原本》的各国文字译本.塔塔里亚在1543年完成了一个意大利文译本;约翰·舒贝尔(Johannes Scheubel)和威廉·霍尔兹曼(Wilhelm Holzmann)分别于1558和1562年将《原本》的主要部分译成了德文;彼埃尔·法卡德尔(Pierre Forcadel)则于1564—1566年间将主要部分译成了法文;1576年罗得里克·卡莫拉诺(Rodrigo Camorano)又完成了前六卷的西班牙语翻译.最有影响的本国语译本是1570年亨利·比林斯利(Henry Billingsley)的英译本.这个将近有1000页的译本包括了《原本》的全部13卷正文,同时还收载了长期以来一直认为是欧几里得的另外三部著作的英译,以及大量古今作者撰写的附录和注解.对于该书的印制,出版商显然是不惜工本.例如第XI卷关于立体几何的讨论就附有一些"可折叠"

的立体图片，粘在相关的书页上，以便读者能实际构作三维图形(图 10.1).

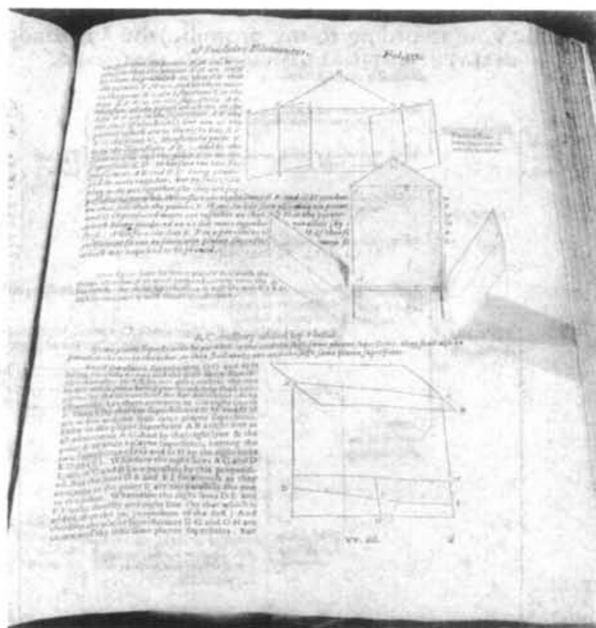


图 10.1　比林斯利翻译的欧几里得《原本》中的一页，内含立体折叠图形.(出处：Library of Congress.)

比林斯利欧氏《原本》中最引人注目的部分是由 16 世纪英国科学家和神秘主义者迪伊(John Dee，1527—1608)所作的“数学序言”(Mathematical Preface).迪伊完全有资格为欧几里得《原本》写序.他通晓各种需要用到数学的领域，并希望使那些打算钻研这本伟大几何著作的人相信它的价值.他详细描述了 30 多个需要数学的不同领域，并指出了它们之间的相互联系，为此还编制了一张他称之为“地图”的框图.迪伊的框图为我们提供了文艺复兴时期“应用数学”的一个概况(图 10.2).

人物小传	迪伊(1527—1608)(John Dee)
	迪伊于 1545 年在剑桥大学获得学士学位，随后赴欧洲大陆与许多数学家一起进行研究，专攻地理学，天文学和占星术，并在巴黎讲授欧几里得几何.返回英国后，他作为一名宫廷占星师为伊丽莎白女王服务.他的著作涉及领域广泛，包括逻辑学，天文学，透视学，凸透镜和占星术，但其晚年倾心于数学中的神秘主义.于是，他研究了如何用各种不同的记号合成特定的图形，据信对这些图形的恰当理解有助于领会现实世界的内在奥秘.像他的一些同时代人一样，迪伊也醉心于词汇的数值研究(gematria)和炼金术.最后，他的神秘主义以及被指控参与“黑魔”妖术活动使其丧失了皇室官职，贫困而终.

J. DEE.

Here haue you (according to my promiſſe) the Groundplat of my MATHEMATICALL Preface: annexed to Euclide (now firſt) publiſhed in our Engliſh tounge. An. 1570. Febr. 3.

Sciences, and Artes Mathematicall, are, either

Principall, which are two, onely,
Arithmetike. Simple, Mixt,
Geometrie. Simple, Mixt,
EVCLIDES ELEMENTES.

Deriuatiue fro the Principalls, of which, some haue

The names of the Principalls: as,
Arithmetike, vulgar: which conſidereth
Arithmetike of moſt vſuall whole Numbers: And of Fractions to them appertaining.
Arithmetike of Proportions.
Arithmetike Circular.
Arithmetike of Radicall Nõbers: Simple, Compound, Mixt: And of their Fractions.
Arithmetike of Coſſike Nõbers: with their Fractions: And the great Arte of Algiebar.

Geometrie, vulgar: which teacheth Measuringe
At hand — All Lengthes. All Plaines: As, Land, Borde, Glaſſe, &c. All Solids: As, Timber, Stone, Veſſels, &c. — Mecometrie. Embadometrie. Stereometrie.
With diſtance from the thing Meaſured, as, — How farre, ... Apomecometrie. How high or deepe, ... Hypſometrie. How broad, ... Platometrie. — Of which are growen the Feates & Artes of Geodeſie. Geographie. Chorographie. Hydrographie. Stratarithmetrie.

Proper names as,
Perſpectiue,
Aſtronomie,
Muſike,
Coſmographie,
Aſtrologie,
Statike,
Anthropographie,
Trochilike,
Helicoſophie,
Pneumatithmie,
Menadrie,
Hypogeiodie,
Hydragogie,
Horometrie,
Zographie,
Architecture,
Nauigation,
Thaumaturgike,
Archemaſtrie,

图 10.2　迪伊为比林斯利翻译的欧几里得《原本》所写序言中的"框图".(出处:Smithsonian Institution Libraries, Photo No.93 – 345.)

作为希腊数学的细心的学生,迪伊在他的序言中指出:"有两类基本的数学对象,即数(Number)和量(Magnitude)."[3] 关于数的科学叫做算术,关于量的科学称为几何.这是数学艺术的两

个基本分支.迪伊认为算术本来只研究整数,但算术家们已经“扩展了数的名称,使其远远超出了那些以单位 1 为最小部分的数.”[4] 各种其它类型的数被引入了,包括普通分数、六十进制(或天文学的)分数和开根数(根数).算术还被扩展成为“方程的算术艺术”,即代数学.然而,序言中的大部分内容是关于“量的科学”即几何学的应用.

迪伊简要地回顾了几何学的历史,以证实几何名称(意为“土地丈量”)的来源,虽然他认为这一名称“对于这样一门高贵而又内容丰富的科学来说是过于低调和狭隘了”.这个名称“得以保留下来,也许含载着对这门科学给普通人带来的最初和最显著的利益的永恒记忆:当土地的边界被毁坏或混淆,当土地作为遗产被转让,或者当公共土地被分配作私有财产时,在所有这些或类似的情形下,有人出于无知或疏忽,错误地丈量、划分了土地,更有甚者,则通过欺骗或使用武力来挑起事端、侵占土地,结果是巨大的损失,动荡、谋杀和战争常常接踵而来,直到由于上帝的慈悲和人类的勤奋,这门关于直线、平面和立体的完美科学被赋予人类自身.”[5] 了解这门科学在它产生之日起就有助于防止战争和维护公正是大有裨益的.

迪伊将几何的应用分成两大类:“普通”几何(vulgar geometry)(包括各种测量的科学,如测体学,即关于立体的测量,和地理学,即关于地图绘制方法的研究)和“方法艺术”(methodical arts),后者“虽不像我们的关于量的基本科学那样具有纯粹、简单和非物质的特性,但仍需使用所谓的基本科学的主要工具、方向和方法.”[6] 这些方法艺术包括透视学、天文学、音乐、占星术、静力学、建筑学、航海学、人体图学(人体几何研究)、圆动学(Trochilike)(圆周运动的研究)、质械学(menadry)(简单机械的研究)和绘图学(关于绘画的研究).本章将考察其中的某些领域,讨论迪伊的分析和 16 世纪及 17 世纪早期一些实践者们的实际工作.

10.1　透视学

根据迪伊所说,“透视是说明所有光线直射、折射和反射的方式与性质的数学艺术”.这种艺术解释为什么“平行的墙在远处互相靠近”,为什么“地板和天花板是平行的,但一个向下,另一个向上,在离你很远的地方互相靠近.”[7] 与透视密切相关的是绘图学,它“讲授并证明怎样用直线来表示视锥与任意平面的交截部分.”[8] 一个画家必须熟知这两项艺术才能“在冬天显示出夏日里欢快的场面,而在夏天展示出冬日里万物萧条的景象.城镇、堡垒、树林、军队、甚至整个王国 …… 所有这些事物他都能驾轻就熟以生动逼真(按任何人的标准)的造型表现出来.”[9]

透视方法虽然在古代就已有人使用,但直到文艺复兴时期,画家们才开始迫切地尝试赋予他们的作品以视觉的深度.起初他们是采用反复试验的方法,但到了 15 世纪,艺术家们开始努力探讨在二维平面上展现三维物体的数学基础.显然,为使绘画具有真实感,离观察者较远的物体应画得较小,问题是一个给定的物体究竟应该被画多小?画家们最终认识到这个问题的解决要依赖几何学.布努雷契(Fillippo Brunelleschi,1377—1446)是第一个认真从事透视几何研究的意大利画家,而阿尔贝蒂(Leon Battista Alberti,1404—1472)(图 10.3)于 1435 年写成了第一本透视学著作,名为《论绘画》(Della Pittura).阿尔贝蒂在书中指出,对一个画家首要的要求是必须掌握几何学.于是他提出了一个几何结果,以说明怎样在画布亦即“画平面”上表现“地平面”上的一组正方形.



图 10.3　意大利邮票中的 L · B · 阿尔贝蒂.

设想画面被不同物体射到画家眼睛的光线穿过,画家眼睛所在位置叫做“定点”(station point),因此画面是一组从眼睛到被画景物的投影截面

(图 10.4). 从定点出发垂直于画面的直线与画面的交点 V 称为"视心"或"中心消失点". 过中心消失点的水平线 AV 叫做"消失线"或"水平线". 使用"消失点"或"消失线"的名称是因为所有垂直于画面的水平线在画中都应被画成在"消失点"处相交. 所有其余的平行水平线组则都将在"消失线"上的某一点处相交.

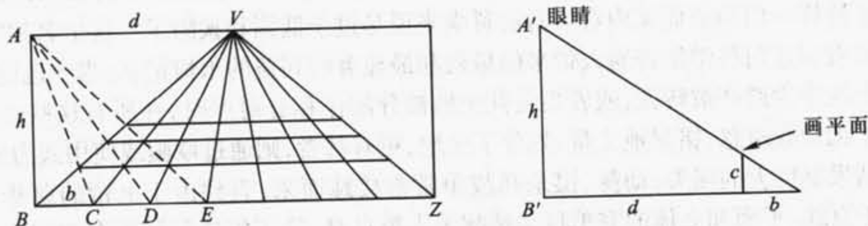


图 10.4　阿尔贝蒂方格地板的透视画法.

为了在画面上表现出由正方形砖块铺成的地面(方砖的对边分别与画面平行和垂直),阿尔贝蒂先在画面和地面的交线("基线")BZ 上作出等距离的点 $B,C,D,E,\cdots$,将这些点与中心消失点 V 相连,得到垂直于画面的一组边. 为了作出平行于基线的另一组边,阿尔贝蒂创造了下述方法:在消失线上作距消失点 V 为 d 的点 A,d 等于画家眼睛与画面的距离,然后作点 A 与点 $B,C,D,E,\cdots$ 的连线,那么过 BV 与 $AC,AD,AE,\cdots$ 的交点且与 BZ 平行的直线就是平行于基线的那一组边.

我们用代数方法证明上述作图法的正确性. 阿尔贝蒂本人并未给出证明. 假设眼睛距地面高度为 h,若地平面上的一条直线平行于基线且在画面之后距基线为 b,那么这条线在画面上应位于基线上方距离为 c 处,这里 c 由如图 10.4 所示的相似三角形导出的比例式 $c:b=h:(d+b)$ 决定. 于是 $c=\dfrac{hb}{d+b}$. 现若取 AB 和基线 BZ 为坐标轴,则连结 B 和 V 的直线方程为 $y=\dfrac{h}{d}x$,而连接 $A(0,h)$ 和 $C(b,0)$ 的直线方程为:$y=-\dfrac{h}{b}x+h$. 于是直线 BV 与 AC 交点的 y - 坐标为 $\dfrac{hb}{d+b}$,这正是所要求的. 容易对所有平行线证明这一结论的正确性.[10]

从 15 世纪至今,这种棋盘结构一直是画家们所使用的"中心透视"系统的核心. 阿尔贝蒂本人没有再讨论更进一步的透视作图,而弗朗西斯卡(Piero della Francesca, 1420—1492)在他作于 1470—1490 年间的著作《绘画透视法》(De prospectiva pingendi)中则详细讨论了如何按中心透视法来画各种各样二维和三维的几何图形. 弗朗西斯卡既是一位艺术家,也是一位出色的数学家,第 9 章中已经提到过他的算盘书中的一些问题. 他的透视法著作中有一幅画显示了画家在准备按中心透视法作画时进行计算的情景.[11]



图 10.5　迪勒自画像.

这一时期的另一位数学艺术家是德国人迪勒(Albrecht Dürer, 1471—1528)(图 10.5). 他在完成自己的两部重要著作之前,曾在意大利度过好几年,专门学习透视法. 第一本著作发表于 1525 年,题为《论线、面、体的尺规测量》(Underweysung der Messung mit Zirckel und Richtscheyt in Linien, Ebnen, und gartzen Corporen),该书也是第一本用德文写成的几何著作.[12] 迪勒必须创造新的德文词汇,来表示包括抽象数学概念在内的科学术语. 他尽量用工匠们代代相传的表达形式. 如"der neue Mondschein"(月芽)表示两圆相交部分,"Gabellinie"(叉形线)意为双曲线,"Eierlinie"(蛋形线)指的是

椭圆.

迪勒相信,在论及《尺规测量》书尾的透视法以前,他需要先教给德国画家们许多关于绘画的几何学准备知识.在他看来,德国画家们在实际技巧和想象力方面毫不比他人逊色,但在理性知识方面却大大落后于意大利人."因为几何学是所有绘画的正确基础,我决定为所有渴望艺术的年青人教授几何的基本原理."[13] 工作是最好的实践.迪勒指出了怎样应用几何原理在画布上表现各种事物(图 10.6).

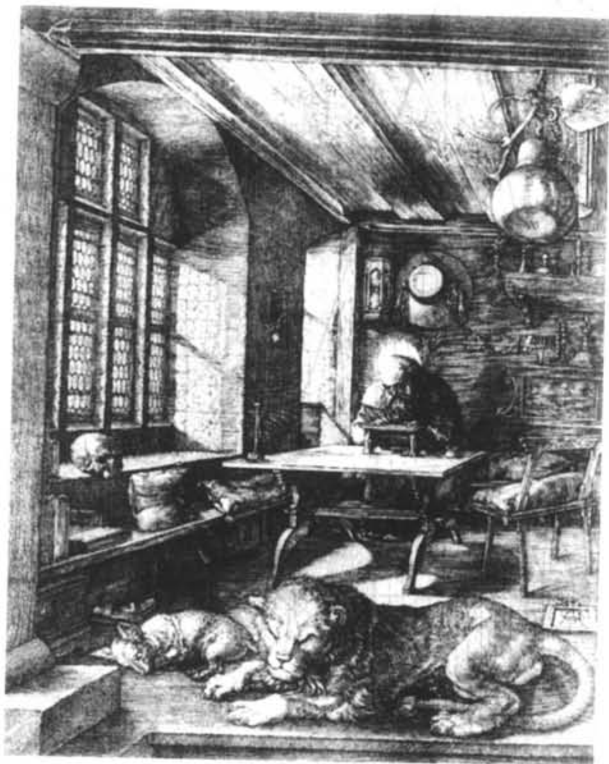


图 10.6 油画《圣·哲罗姆(St. Jerome)在书斋》.迪勒在这里图示了他的透视理论的应用.(出处:The Nelson-Atkins Museum of Art, Kansas City, Missouri, 58 - 70/21, 菲泽尔(Robert B. Fizzell)赠.)

迪勒《尺规测量》四卷中的第一卷处理空间曲线的表示问题.他的思想是将曲线同时投影到 yz 面和 xy 面上以确定其性质.不幸的是,这并非总是一件简单的事情.以他的椭圆作图为例.从椭圆作为正圆锥面截面的定义出发(图 10.7),迪勒首先将圆锥及其截面投影到 yz 平面.将表示椭圆直径的线段 fg 分为12等份,过每一分点同时作垂直和水平的直线.在11个分点的每一点 i 处,水平线表示圆锥被一水平面截得的圆截面 C_i 直径的一部分.圆截面与椭圆的两交点相互对称并位于椭圆直径的两端,因此确定椭圆截面在该处的直径长为 w_i.圆锥投影到 xy 面上得到一列同心圆 C_i.每一条垂线延伸为相应圆的弦,弦长为 w_i.于是迪勒有了一个椭圆的粗投影.然而,这个椭圆的轮廓关于其短轴并不对称,因为投影不是沿垂直于椭圆本身的方向进行的.但当迪勒试图从投影图来画原椭圆时,他只是简单地将表示椭圆轴的线段转换为垂直线 fg,并在同样的那些点 i 处将它分段,然

后在每个分点上作相应的长为 w_i 的水平线段,再通过这些线段的端点描画出曲线的略图.迪勒的画法是错误的,因为画出的曲线下方比上方要宽.原因可能是迪勒没有认识到椭圆应该是关于其短轴对称的图形,圆锥面的中心线并不通过椭圆的中心,而所有的圆都是围绕它的投影画出的.尽管通过分析可以证明 $w_i = w_{12-i}(i = 1,2,3,4,5)$,迪勒可能认为椭圆实际上是蛋形的 —— 他确实管它叫 Eierlinie—— 因为锥面本身由上而下变宽.[14]

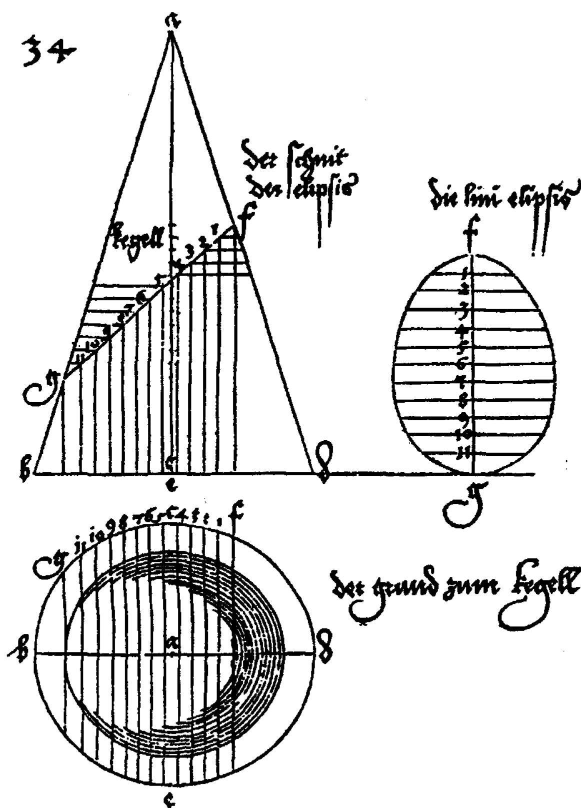


图 10.7 迪勒的射影椭圆作图.

在描述了其它一些平面曲线的投影图示之后,迪勒在《尺规测量》第二卷中继续描述各种正多边形的作图法,既有借助经典工具 —— 直尺和圆规的精确作图,也有根据艺人传统的近似画法.因此这本著作(其德文本出版几年后又出版了拉丁文本)起到了双重作用,即一方面向艺人们介绍希腊几何经典,一方面使职业数学家熟悉工场中的实用几何知识.该书的第三卷是纯粹实用性的,说明几何学如何被应用于诸如建筑及印刷等各个领域.迪勒在这里建议了若干新型的圆柱和屋顶设计,提出了罗马体和哥特体铅字的准确构造法.迪勒在该书最后一卷中回到了更为经典的问题,论述了三维体的几何.特别是,他提出了一种通过折纸构作五种正多面体的方法,以及类似的构作某种半正多面体的方法,这种方法在今天的教科书中仍能找到.他还给出了包括倍立方问题在内的其它一些问题的作图法,最后以关于这些立体图形透视画法的基本规则结束全书.

迪勒最富创造性的几何著作是他的《人体比例四论》(Vier Bücher von menschilcher Proportion),1528 年出版.该书内容包括"人体图学" 的研究,即"关于构成完美的人体的各不同部分的数量、长度、重量、形状、位置和颜色的描述,以及关于所涉人体的特定部分的对称性、形状、轻重、特征和局

部运动的某些知识".[15] 迪勒列出了对不同尺寸的人体的一系列测量数据，并讨论了人体各个部分特别是头部在所有三个坐标平面上的投影. 该书四卷中的最后一卷处理运动的人体. 迪勒这部著作，像达·芬奇(Leonardo da Vinci, 1452—1519) 的一些早期著作一样，证明了在画布上正确描绘人体的背后有许多几何学. 以往的处理如今被精心制定的数学理论所代替，这种数学理论注定将在随后的几个世纪中产生广泛的影响.

10.2 地理和航海

迪伊所讨论的两个与数学相关的、对16世纪的世界来说也极端重要的领域是地理和航海."航海术讨论怎样导航(处于航道上的) 两指定地点之间的船只，使其通过最短的有效路径，沿最恰当的方向并在最短的时间内抵达目的；以及在遭遇各种风暴和自然障碍时，怎样使用最佳可能方式来恢复最初的航向."[16] 在15和16世纪，欧洲人正在探险世界的其余部分，航海术至关重要. 能够较好运用这方面的新技术的国家，在探寻新殖民地及其拥有的自然资源方面具有巨大的优势.

海上航行的主要问题是确定船只在任意给定时刻的经纬度. 纬度的确定并不十分困难. 在北半球，一地的纬度等于北天极的高度，而北天极的位置可用北极星近似标记. 这一高度的良好近似值可以通过北极星的高度而得到(虽然在15世纪时北极星与北天极有3.5°的偏差，因而需要作出适当的调节). 在赤道附近或赤道以南，求纬度的一种替换方法是观测太阳. 正如第4章所述，在当地正午太阳的天顶距等于纬度减去太阳的赤纬. 15世纪的航海家已有一年中每天的精确赤纬表，因此他们只需要观测出太阳在正午的高度. 当然，这个高度是当日最高的高度，可以通过求出某一标杆的最短影长来确定.

经度的确定要困难得多. 由于15°经度相当于1小时，故要知道两地的经度之差相当于要知道两地的当地时差. 从理论上讲，如果把一个时钟按照已知经度的某地时间来设置，并能确定在另一地点当地的正午在该钟上何时出现，其时差便能确定出当前地点的经度. 另外，也可以比较某一天文现象(如月食) 在已知经度的某地所发生的时间与它在当前位置发生的当地时间. 不幸，就当时对月球运动的知识或计时设备的精度而言，这些方法都不可行. 那时的时钟不够精确，特别是在摇晃的甲板上，加之在海中船上的温度变化不定. 哥伦布于1494年第二次航行到美洲，当他试图用月食确定经度时，其误差大约是18°. 晚至1707年，四艘英国军舰因舰队司令和导航员对船队经度判断错误而在英国西南端附近的锡利群岛(Scilly Isles) 触礁，死亡2000人. 英国政府后来对精确确定大海经度的方法悬赏20 000英磅. 这笔钱(至少绝大部分) 最终付给了一位英国钟表匠，约翰·哈里逊(John Harrison, 1693—1776)，他的一系列的精度愈来愈高的计时器在陆地和大海上经受了大量的考验，并在詹姆斯·库克船长航行南太平洋时赢得了他的赞扬(图10.8). 大约在同时，经过皇家格林威治天文台近一个世纪的详密观测，其精度足以满足确定经度之需的月亮数据表开始面世了.[17]



图10.8 英国邮票中哈里逊最后的计时器.

知道确定在大海中位置的困难之后，就不应该吃惊于水手们为什么经常使用"猜估"而非数理天文学的方法. 虽然学者们已清楚大圆路径是两点之间的最短距离，水手们一般却宁愿尽快航行到目的地所在的纬线，然后再向正东或正西前进直到靠岸. 然而，无论采取什么航行方式，水手们都需要精确的地图. 迪伊把制作这些地图称之为地理学："地理学教我们以各种形式(如球面，平面，或其它) 模拟自然和真实来描绘地表

上诸类事物的具体位置，如城市，村镇，要塞，城堡，山脉，河流，港口，森林，小道，及其它，…… 并表示成最便于我们观察的形式”.[18]

地图的制作源于古代. 托勒密在他的《地理学》中已经分析了将圆形地球画到平面纸上的一些问题，显示出所居住世界的那些已知地点的经纬度，他制作了大约 26 个区域地图和一张世界地图. 为了构画他的地图，他需要使用某种形式的投影，即某种构造把球形地表的一部分对应到一块平面纸上的函数的方法. 可以想象，托勒密希望制成的地图能尽可能准确地表示所描绘的那些陆地的形状. 在任何情况下，投影取决于由经纬线(即一般所称的子午线和纬圈)组成的坐标方格. 对于他的区域地图，托勒密简单地使用矩形方格来表示这些线. 但由于子午线的间隔大小依赖于纬度，他选择了一种双向的标尺，以使其近似对应于地图中间纬线上 1° 的经线长度与 1° 的纬线长度之比. 这个比率等于 $MN : AB$(因为赤道上 1° 的经线长度与 1° 的纬线长度相等)，进而等于 $NP : BC$，$NP : NC$，最终为 $\cos\phi$，其中 ϕ 是中间纬度(图 10.9). 例如，托勒密的欧洲地图是从北纬 42° 到北纬 54°，所以，给定的比率应该近似于 $\cos 48° = 0.6691$ 或 $2 : 3$.

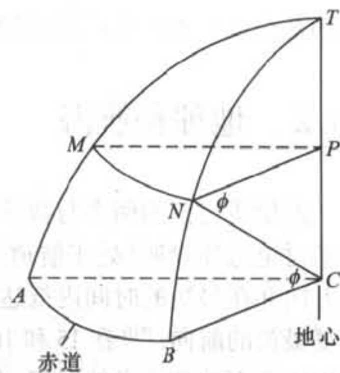


图 10.9　在纬度 ϕ 处 1° 的经线长与赤道上 1° 的经线长之间的关系.

对于他的世界地图(其实仅包括了他所计算过的 180° 的经线范围，即从直布罗陀海峡到中国)，托勒密曾用过两种不同的方法. 在第一种方法中，纬线被表示成以北极为中心的同心圆，子午线则是从北极出发的直线，而北极未能画进图中. 托勒密意识到这种投影不能保持经线度数与纬线度数



图 10.10　托勒密《地理学》中的世界地图，1552 年[瑞士]巴塞尔版本.(出处：Smithsonian Institution Libraries, Photo No. 90 – 15 779.)

之间的适当比率,除非在某特殊纬线(他取其为罗得岛所在的36°纬线)周围的一个较小区域内.因此,对于地表的较大区域的投影来说,某种程度的失真是不可避免的.后来,托勒密通过把子午线也修改为圆弧而发明了一种更自然直观的地图.这种地图对于所选择的三条纬线(圆弧通过它们画出)上的距离,给出了正确的表示,但在远离地图中心的部分仍有失真(图10.10).

由于不可能在纸平面上绘制出完全准确的地图,作图者一般需要对所要求的投影特性作出选择.作图者可以选择保留面积,或形状,或方向,或距离.所要表示的地表的区域越大,越难以在地图上同时保留其中的几个特性,即使近似也不行.一般来说,在文艺复兴早期水手们所用的地图是按照一种不同的标准——即容易绘制——来制作的.这些"平面航线图"对经纬线采用矩形方格,并在整个图上采用相同的比例尺度.由于在地图的所有纬度上两条子午线间的距离都相同,而实际距离却依赖于纬度的余弦,故这些地图的形状呈现出在水平方向被伸长的样子.因此,形状未能得到保留,而对于水手来说更重要的是,恒定的罗盘方位线(亦称"罗经方位"[rhumbs])也未被画成直线.当这样的地图表示较小的区域时,罗盘方位线还相当直,通常画出来表示6个或8个罗经航向.但当漫长的海上航行变得越来越普遍时,制图方法的改进就势在必行了.

早期试图应用数学改进地图制作方法的人物之一是努涅斯(Pedro Nunes).在1537年的著作《球面论》(Tratado da sphera)中,他发现球面上的罗盘方位线,亦即现今所说的斜驶线(loxodrome),是一条以极点为终点的螺旋线.然而,携带地球仪航行是不方便的,因为它们不可能被做得足够大.因此,努涅斯试图发明一种地图,使斜驶线成为直线.然而,为了精确起见,子午线必须收敛于两极.虽然努涅斯能够设计出一种仪器使水手可以测定沿着每一度纬线的距离是多少英里,但他无法解决他所提出的这个问题.

直到1569年,努涅斯的问题才被墨卡托(Gerard Mercator,1512—1594)从略微不同的观点所解决(图10.11),墨卡托使用了一种新的"墨卡托投影法".在这种地图上纬线和子午线都被表示成直线.为了补偿子午线的"不准确"间隔,墨卡托增加了两极方向纬线的间隔.他声称在他的地图上罗盘方位线是直的,航海家只需简单地将一根直尺放在起点和终点上就能确定所要遵循的罗经航向.墨卡托没有解释在增加纬线之间的距离时他所遵循的数学原理,因而有人认为他仅仅依赖于推测.直到E·赖特(Edward Wright,1561—1615)的著作《论航海中的失误》(On Certain Errors in Navigation,1599)正式出版,才论证了墨卡托的地图方案.



图10.11 比利时邮票中的墨卡托.

我们前面已经知道,在纬度 ϕ 处1°的经线长度与赤道上1°的经线长度之比等于 $\cos\phi$.如果经线是直线,那么在纬度 ϕ 处,它们之间距离的伸长因子就应该是 $\sec\phi$.由于在这种地图上斜驶线是直的,其垂直距离的伸长因子也应该相同.因为沿着一条子午线 $\sec\phi$ 在每一点都发生变化,故应该相对于纬度的每一微小变化来考虑伸长因子.如果我们用 $D(\phi)$ 表示地图上赤道与纬度 ϕ 处的纬线之间的距离,则由 ϕ 的微小改变量 $\mathrm{d}\phi$ 引起的 $D(\phi)$ 的改变量 $\mathrm{d}D$ 可以这样确定:$\mathrm{d}D=\sec\phi\mathrm{d}\phi$.因为同样的因子也适合于水平方向,故地球仪上的任何"小"区域可以在地图上用相同形状的"小"区域表示出来.地球仪上某条线与子午线的交角将转化为地图上相同大小的角,且斜驶线将是直的.由此可知,若用现代术语表示,在纬度 ϕ 处赤道与纬线之间的地图距离可以由下式确定

$$D(\phi)=\int_0^{\phi}\sec\phi\mathrm{d}\phi, \tag{10.1}$$

其中地球仪的半径取为1.当然,赖特没有使用积分形式.他将 $\mathrm{d}\phi$ 取作为1′的角,通过添加乘积 $\sec\phi\mathrm{d}\phi$,计算出一张纬度75°以内的"子午线数据表".事实上 $D(\phi)$ 可用微积分求知,为

$$\ln(\sec\phi+\tan\phi)\quad 或\quad \ln\left(\tan\left(\frac{\phi}{2}+\frac{\pi}{4}\right)\right).$$[19]

迪伊在一次去欧洲大陆的旅途中会见了墨卡托.他返回时带有几个墨卡托的地球仪,可能还与赖特讨论过墨卡托投影的数学依据.迪伊曾致力于"模拟自然"的地图制作.墨卡托的地图,虽能很好地适用于航海,但不幸在远离赤道的地区却并不能"模拟自然".在高纬度地区的区域相对地变得越来越大.这种地图的流行导致一代一代的学生相信,譬如,格陵兰比南美洲还大.然而,其使用的简便性使它成为欧洲探险时代的主要的航海图.

10.3　天文学和三角学

根据迪伊所说,"天文学是一门数学的艺术,它论证在过去、现在和将来的任何时刻,行星和恒星的距离,大小,运动,出没,及反常现象.借助于这种艺术,我们可以证实任何可见恒星或行星相对于地球而言的大小".[20] 因此,天文学的目的是预测天体的运动,确定它们的大小和距离.一门相关的艺术是宇宙学,它是关于"天地万物的完美描述,…… 以及必要的相互对照."[21] 正如迪伊进一步所述,宇宙学解释天地现象的关系,使人们能确定"太阳的升落,昼夜的长短,…… 以及其它许许多多有趣而必要的应用."

10.3.1　雷格蒙塔努斯

由于文艺复兴时期的天文学和宇宙学同较早时期一样,完全依赖于三角学,我们首先来讨论欧洲的第一部"纯"三角学课本——《论各种三角形》(De Triangulis Omnimodis),作者原名缪勒(Johannes Müller,1436—1476),因其生于柯尼斯堡(Königsberg in Lower Franconia)附近,通常被叫做雷格蒙塔努斯(Regiomontanus).(《论各种三角形》大约写于 1463 年,但 70 年之后才印刷出版.)

雷格蒙塔努斯曾直接从希腊文翻译了托勒密的《天文大成》,在完成翻译之后,他意识到需要有一本简洁系统的著作来论述平面和球面三角形边角关系的法则,这势必会强化对托勒密的那些表面上很特别的处理方法的理解.他视其为学习《天文大成》的前提:"若想研究天体运动,探索其奥秘,必须先阅读这些关于三角形的定理.这些知识是通往整个天文学和某些几何问题的必由之路."[22]

在《论各种三角形》中,雷格蒙塔努斯按照几何模式用定义和公理开头来精心组织材料.他用公理、欧氏《几何原本》的结论、或书中前面的结论来证明每一个定理.大多数定理都有附图,还有解释性的具体例题.雷格蒙塔努斯的三角学是以弧的正弦为基础,正弦的定义是二倍弧的半弦,但他也指出,可以把正弦视为依赖于相应的圆心角.像他的欧洲先辈们一样,他没有使用正切函数,虽然他肯定知道那些当时在欧洲已能见到、多半源于伊斯兰天文著作的正切表.这些表作为具有特定尺寸的日晷指针的影长表,不仅出现在当时的课本中,还出现在许多星盘上以使天文计算更为方便.雷格蒙塔努斯也许认为它们对于他的理论性阐述没有必要,而未将其纳入自己的著作.但在 1467 年完成的另一本天文表汇编中,他确实收录了一个正切表,并称之为"富有成效的表"(tabula fecunda).总之,雷格蒙塔努斯仅用正弦表便能解决三角学的一切标准问题,他在书后附了一张基于半径(或全正弦)60 000 的扩充了的正弦表.

雷格蒙塔努斯著作的前半部论述平面三角形,后半部处理球面三角形.在他的结论中含有各种解三角形的方法.从概念上讲,他的方法并没有特别之处,但和欧洲早期的作者不同,他总能提供出

清晰准确的例题.例如,定理 I - 27 论述如果给定两边长,如何确定直角三角形的角;定理 I - 29 论述如果给定一个锐角和一边长,如何确定直角三角形的未知边.在两种情况下,雷格蒙塔努斯都使用他的正弦表.他针对后一定理的例题是,假定一个锐角为36°且斜边为20.由于另一角为54°,当斜边为60 000时,另外两边将分别为35 267和48 541.应用比例法,雷格蒙塔努斯便算出,当斜边为20时,另两边分别为 $11\frac{3}{4}$ 和 $16\frac{11}{60}$.定理 I - 49论述如果已知两边及其夹角,如何解任意三角形.假设已知两边 AB 和 BC,及其夹角 ABC,雷格蒙塔努斯采用了与莱维本热尔松一样的方法(图10.12).从 A 点作 BC 或 BC 延长线的垂线 AD.在直角三角形 ABD 中,一个锐角和一边长度已知.由定理 I - 29,其余边和角便能被算出.直角三角形 ADC 的两边长现在为已知,由定理 I - 27,另一边和角也能算出.

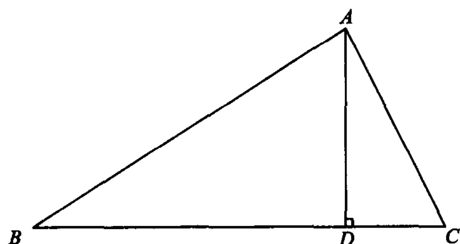


图 10.12 《论各种三角形》:定理 I - 49.

对于所谓的二义情况,即已知两边及其一边的对角,雷格蒙塔努斯通过提供几种可能性改进了莱维本热尔松的结果.他首先通过从 A 点作 AC 延长线的垂线 AD,讨论了已知边 AB 的对角 ACB 被给定为钝角、且边 AC 也为已知的情况(图 10.13).这个三角形可以像前面的问题一样解之.然而,在论述给定角是锐角的情况时,他指出"所给条件不足以求出另一边及其它角."[23] 对于锐角 ABC 的对边 AC 而言,有两种可能的三角形,在第一种情形中,边 AB 的对角为锐角,在第二种情形则为钝角.他论述了在每一种情况下,如何求未知边和角,但未能指出也许无解的可能性,究其原因,他可能以为那个欲求其未知部分的具体三角形肯定存在.

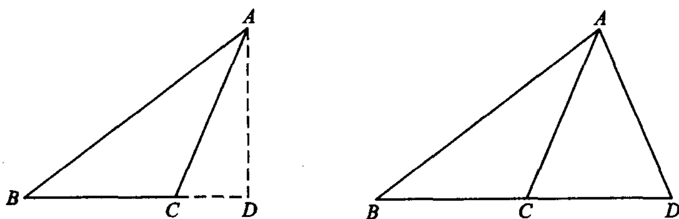


图 10.13 《论各种三角形》:二义情况.

在定理 II - 1 中,雷格蒙塔努斯证明了正弦法则:"在每一直线三角形中,一边与另一边之比等于一边对角的正弦与另一边对角的正弦之比."[24] 由于他的正弦是给定半径的圆中的直线,他对三角形 ABG 的正弦定理的证明就需要分别以 B 和 G 为圆心,以相等的直线 BD 和 GA 为半径画出两个圆(图 10.14).再从 A 和 D 分别作 BG 的垂线,依次相交于 K 和 H,雷格蒙塔努斯然后指出,应用相等半径的圆可知,DH 是 $\angle ABG$ 的正弦,而 AK 是 $\angle AGB$ 的正弦.因为 $BD = GA$,$\angle ABG$ 的对边为 GA,且 $\angle AGB$ 的对边为 AB,所以由三角形 ABK 和 DBH 的相似性可得出要证明的结论.雷格蒙塔努斯随后便用这个结论重新来解已知两边及其一边之对角的三角形.

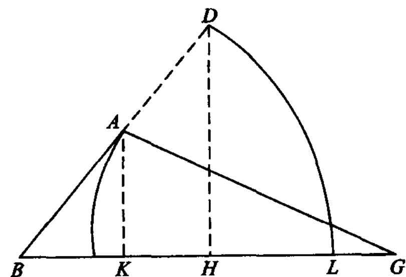


图 10.14 正弦定理的证明.

在《论各种三角形》卷 II 的剩余部分,雷格蒙塔努斯论

述如果给定某些条件，诸如两边的比率或者某一顶点到对边的垂线长度，如何确定三角形的各种量.其中有两个定理，他没有采用几何论证，而是借助于代数推理(他称之为"物和平方之术")，因为他声称对其结果无法提供"几何"证明.这样，给定某一三角形的底边 $BG = 20$，垂线 $AD = 5$，比率 $AB:AG = 3:5$，为求两边 AB 和 AC，雷格蒙塔努斯设线段 $DE = BD$，且为代数上简便起见设未知线段 $EG = 2x$(图 10.15).那么 $BE = 20-2x$，$BD = 10-x$，且 $DG = 10+x$.因为 $AB^2 = BD^2 + AD^2$ 和 $AG^2 = DG^2 + AD^2$，且比率 $AB^2:AG^2 = 9:25$，雷格蒙塔努斯得出

$$\frac{(10-x)^2+25}{(10+x)^2+25} = \frac{9}{25}.$$

这个方程容易化简为 $16x^2 + 2000 = 680x$.雷格蒙塔努斯在此处停止了求解，仅指出"剩下要做的，按代数法则进行."[25]

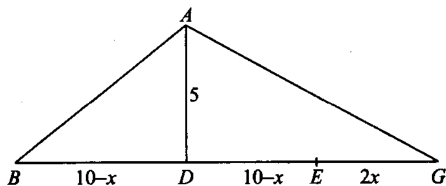


图 10.15　雷格蒙塔努斯所用的代数方法.

《论各种三角形》卷 III 对球面几何学提供了一个基本的介绍，特别是包含了许多大圆的结果.本卷是其后两卷球面三角学的标准论述的预备知识，雷格蒙塔努斯在这两卷中包括了直角和任意球面三角形的正弦法则，以及解这种三角形的各种方法.卡尔丹评论说，大部分材料似乎源于贾比·伊本·艾夫拉赫(Jābir ibn Aflah).事实上，对雷格蒙塔努斯的定理与其西班牙前辈的相应结果的仔细比较表明，在许多情形中不仅证明相同，甚至使用的图标文字也相同.[26] 虽说如此，雷格蒙塔努斯著作的最后两卷包含了比贾比更多的材料，尽管他引用了伊斯兰(及其它地方)的原著而未加声明，但他的著作确实提供了欧洲最早的关于平面和球面三角学的综合性系统处理.有趣的是，与以往的三角学著作一样，该书未曾包含应用平面三角学的结果来解日常三角形的具体例子.

在雷格蒙塔努斯《论各种三角形》之后，16 世纪的后三分之二期间出现了约二十本论三角学的其它著作，其中大多与他的著作雷同.[27] 有些作者改进了他的正弦表并编印了所有其它的三角函数表.一般来说，与正弦一样，它们都被定义为某一固定半径的圆中依赖于某一给定弧的直线的长度.圆的半径通常取 10^n 或 6×10^n. n 的值在世纪末时趋向于更大，这是由于天文计算日益需要更高的精确度，较大的半径可使所有的值取为整数，而十进制小数仍未被使用.然而，雷蒂库斯(George Joachim Rheticus，1514—1574)在他的著作中，把其中一边固定为某一较大数值，直接用直角三角形的角定义三角函数.因而，雷蒂库斯在固定斜边之后，称正弦为"垂线"，余弦为"底线".其他作者也给出三角函数的另外名称.首次使用现代术语"正切(tangent)"和"正割(secant)"的作者是芬克(Thomas Finck，1561—1656)，他于 1583 年撰写了他的著作《圆的几何》(Geometria rotundi libra XIV).他称三种余函数为"补弦(sine complement)"，"补切(tangent complement)"，和"补割(secant complement)".这些三角学课本的大多数都给出各种数字例题来描述解平面和球面三角形的方法，但直到 1595 年皮蒂斯楚斯(Bartholomew Pitiscus，1561—1613)的著作出现之前，这些课本中并没有出现任何明显涉及真正的平面三角形求解问题.事实上，是皮蒂斯楚斯发明了"三角学(trigonometry)"这一术语.他给自己的书起名为《三角学，或三角形测量之书》(Trigonometriae sive, de dimensione triangulis, Liber).

皮蒂斯楚斯试图在课本中阐述如何测量三角形，在关于测高术的附录中，他给出了确定塔高 BC 的三角学方法.在图 10.16 中，用象限仪测知 $\angle AKM = \angle ABC = 60°20'$.从观测者到塔脚的距离测定为 200 英尺.皮蒂斯楚斯建立比例式

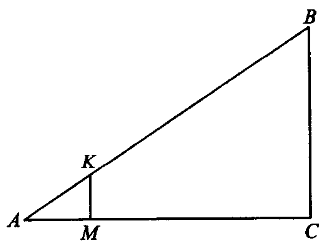


图 10.16　塔高的测量.

$$\sin 60°20' : AC = \sin 29°40' : BC,$$

然后算出 $BC = 113\frac{80\,204}{86\,892}$,或近似地为 114 英尺.这种算法用到皮蒂斯楚斯的正弦表,该表取半径为 100 000.他还用他的正切表给出另一算法,其中要用到比例式 $AC:100\,000 = BC:\tan 29°40'$.皮蒂斯楚斯的方法与现今的主要区别在于,由于其三角值是一特定圆的某些线段的长度,他必须经常调整自己的计算.现今通用的三角比便呼之欲出了.

10.3.2 哥白尼和日心体系

即使没有三角比,以雷格蒙塔努斯的著作为典范的 15 世纪三角学仍为解决当时的天文学问题提供了必要的数学工具.这些问题中有一些在雷格蒙塔努斯的《天文大成》译本中有所讨论,它们涉及托勒密体系的基础,后者在当时仍然是被普遍接受的宇宙结构观.若干世纪以来,伊斯兰和犹太天文学家已经注意到托勒密的预告与他们自己的观测之间存在偏差,并对托勒密的理论作了许多细节上的修正.但在文艺复兴之初,基督教的宇宙观仍然是以亚里士多德和托勒密的观点为基础,坚持认为宇宙由若干层以地球为中心的天球组成,正是这些镶嵌着行星的天球的旋转引起各种天文现象.这种模型的各种附加的修正,诸如本轮和偏心圆,全都以某种方式被附着于各个天球之中.这种基本的宇宙观很容易从但丁的《神曲》(Divine Comedy,1328)中看出,其中描述了诗人越过每个行星的天球到上帝神座所在的不动天球的旅行.

人物小传	哥白尼(1473—1543)(Nicolaus Copernicus)
	哥白尼(图 10.17)生于东普鲁士托伦(Torun)的一个富裕商人家庭,18 岁时被送入克拉科夫(Cracow)大学学习.离开克拉科夫时,由于其伯父——埃姆兰地方主教的影响,他被任命为一名牧师.这样,他不仅能拿到薪俸,而且获准去意大利的波伦亚和帕多瓦学习了十年.返乡后,他在埃姆兰作为弗劳恩堡大教堂的一名神甫度过了余生.他的工作并不费神,故有充分时间专注于自己的天文学研究,并于 1530 年完成了他的《天体运动论》手稿.然而,他并不想立即出版这部著作.早在 1514 年左右,他就已写出了他的太阳系学说的梗概《要释》(The Commentariolus),其手稿曾在学者们中间广为传阅.但直到 1539 年维腾贝格大学的数学教授雷蒂库斯到弗劳恩堡亲自了解了哥白尼体系,哥白尼才终于被说服同意出版自己的著作.



图 10.17 匈牙利邮票中的哥白尼和他的太阳系.

到 15 世纪时,天文学家已很难全部接受托勒密的体系.雷格蒙塔努斯所指出的偏差之一是,托勒密的月球理论需要对观察值作远大于实际所需的修正.更重要的,由于微小偏差已累积许多世

纪,天文学家发现多数情况下托勒密关于行星位置或月食的预报均有很大误差.欧洲探险家在环球航行时需要改进航海技术,而这只有通过精确的天文表才可能实现.另外,通过这些探险,欧洲人发现了许多前所未闻的地域,这使他们认识到托勒密的《地理学》也充满错误.凡此种种均预示,托勒密的天文学基础可能是错误的.

到文艺复兴早期,天主教会也认识到始于古罗马帝国的儒略历存在严重缺陷.特别是,由于实际回归年比儒略历所依据的 $365\frac{1}{4}$ 天少 $11\frac{1}{4}$ 分,这方面误差的累积影响到日历月份与季节之间的关系.例如,按照教会法律,复活节定为每年春分月圆后的第一个星期日.春分一向被推算为 3 月 21 日,但到 16 世纪时,真正的春分已提早到 3 月 11 日.若再不纠正,复活节将会在夏天而不是在春天到来.然而,当教会宣布想改革历法时,天文学家却认为当时的天文观测还不够完善,历法改革尚缺乏精确的数学基础.

这些拒绝接受邀请参加历法改革的天文学家之一是哥白尼(Nicolaus Copernicus,1473—1543),他详细地研究了托勒密体系,并已认识到地心说千疮百孔、无法补救."[天文学家们]一直未能从他们的假设中发现或推演出根本性的东西,即宇宙的构成形式及其各部分的明确对称性.犹如某人在一幅油画中画上栩栩如生的手、脚、头、及其它肢体,虽然画得很好,但却不是以同一个身体为原型,至少互不匹配,因而所创作出的是一个怪物,而不是人."[28] 为了重构这幅"油画"并消除怪物,哥白尼决定阅读古人的观点以寻求是否有人提出过不同于地心说的宇宙体系.当发现某些古希腊哲学家曾提议过地动日心体系时,哥白尼便努力探索在此假定下改革宇宙体系的重要意义:"因此,在地球运动的假设下,经过长期深入的研究,…… 我终于发现,如果把行星的运动看作是和地球一样的圆周运动且按照各自的公转来推算,则不仅天象与结论相符,而且所有星体和天球的分布、大小以及整个天穹彼此密切联系在一起,任何部分的分离都将造成其它部分乃至整个宇宙的混乱."[29]

哥白尼详述其宇宙体系的代表作是《天体运行论》(De Revolutionibus Orbium Coelestium),该书凝聚着他毕生的心血,但直到 1543 年他去世时才出版.《天体运行论》破天荒第一次提出了以地动假设为基础的天体运行的数学描述,因为正如他在前言中所说:"数学是为数学家而写的."[30] 该书在风格上模仿托勒密的《天文大成》,是一本高度专业化的著作,其中作者用仔细的数学计算,以日心假设为基础,以哥白尼本人及其前辈的观测结果为依据,来描述月亮及行星的轨道,说明这些轨道如何通过它们在天空的观测位置反映出来.哥白尼在《天体运行论》第一卷中简要概述了他的理论体系,并给出了一张简图,呈现出太阳位于七个同心天球的中心,其中六个分别代表包括地球在内的六大行星,而处于最外层的是恒星天球(图 10.18).

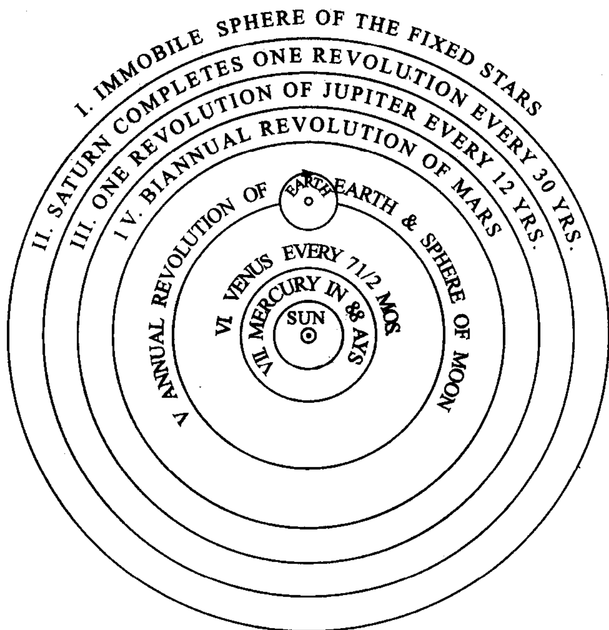


图 10.18　哥白尼的日心体系.

应该强调指出哥白尼也像他的先辈

们一样,将宇宙体系想象为一系列嵌有行星的层层套叠的天球,而不是行星在其中作圆周运动的空虚空间. 哥白尼没有用以解释行星如何保持其轨道的物理学. 而天球对于哥白尼就像对亚里士多德一样,具有一种天然的运动而毋需其它任何物理基础:“天体的运动是一种圆周式旋转,其形状由运动本身呈现出来,圆是最简单的图形,这里既没有起点,也没有终点,相互不可区分,因为旋转总是在同一位置进行.”[31] 事实上,哥白尼改革托勒密体系的目的之一就是要恢复一种经典的天文学原则,即所有的天体运行必由围绕圆心的匀速圆周运动合成. 哥白尼相信托勒密使用等距点(equant)来说明行星运动的某些问题是违背了这一原则(图 10.19). 所谓等距点是指行星轨道上一点,行星均轮中心半径向量围绕着它作匀速转动,而在这种情况下匀速运动并不是相对于均轮运行的圆的中心而言. 在马拉加(Maragha)以纳西尔·丁(Nasir al-Din al-Tūsī)为首的伊斯兰天文学家们也曾为这个问题而困顿. 虽然还不清楚哥白尼是怎样获悉这些伊斯兰人的工作的,但他在自己的著作中改写了他们的研究结果. 当然,伊斯兰天文学家未能像哥白尼那样迈出关键的一步:向地球的中心地位挑战.

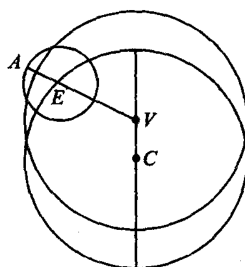


图 10.19　托勒密的等距点(equant):行星 A 沿本轮以 E 为中心运行;E 沿均轮以 C 为中心而动,其中向径 VE 以等距点 V 为心而均匀运动.

哥白尼没有 —— 也不可能 —— 对地球绕其轴的每日自转和绕太阳的每年公转提供任何真正的证据. 对于前一种运动,他简单地推理说,假定相对较小的地球而不是巨大的恒星天球转动要更为合理. 对于第二种运动,他的论证实质上是说,通过把行星运动的一部分归因于地球自身的公转更容易理解行星的定性行为. 这样,行星的逆行现象可用地球及行星的组合轨道运动而不是用本轮来解释(图 10.20). 行星与地球距离的视变化也更容易用这两个轨道来解释. 对地球绕太阳运动的一种反对意见是认为这样会引起恒星在每年不同时间出现在不同的地方(所谓的周年视差),对此哥白尼的回答是:因为地球的轨道半径相比恒星天球的半径小的多,故而观测不到这种视差. 因此,哥白尼学说的作用之一就是大大扩大了天文学家们对宇宙大小的估计.

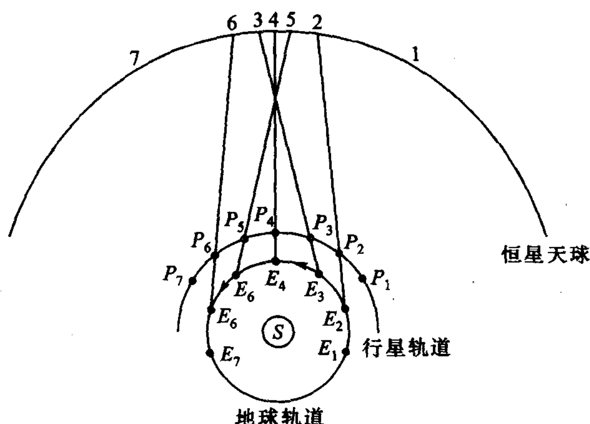


图 10.20　地球轨道之外的行星的逆行现象. 在恒星天球上所观察到的行星位置依次标为 1,2,3,4,5,6,7. 逆行发生在 3 和 5 之间.

在对他的新体系作了基本介绍之后,哥白尼仿效其先师托勒密,给出了一个为解决天体运动中提出的数学问题所必需的平面和球面三角学知识概要. 尽管雷格蒙塔努斯的著作出现后欧洲的三角学技巧已有长足的进步,但哥白尼的论著在这方面仍保留着公元 2 世纪时天文学家著作的特点,甚至包括弦的使用. 然而,他利用自托勒密时代以来 1400 余年的成就作出两点改进. 第一,他用

100 000 而非古代的 60 作为圆的半径(因为此时阿拉伯数字已普及). 第二,他的表给出的不是弧本身所对的弦,而是二倍弧的半弦,"因为这种半弦在推理和计算中比整弦使用更为频繁."[32] 哥白尼没有使用当时已有的术语"正弦"(sine). 在解平面三角形的方法概要中,他也没有提及正弦法则. 他的基本解法是作适当的垂线,然后讨论直角三角形.

哥白尼在论述平面三角形之后给出了球面三角形的若干结论. 在这里他也是只使用弦(或半弦). 他不仅没有给出球面三角形的正弦法则,而且也没有给出从其它三角关系推出的门纳劳斯结论的简化形式. 然而,他关于如何解球面三角形的讨论是全面的,并且对于这部天文著作来说是够用的.

在该书的其余各卷,哥白尼使用他的新日心模型以及古代和当代的观测来计算月球和行星轨道的基本参数. 正是在阅读后面这几卷时,读者可以看出将宇宙中心从地球移开并未能使托勒密的图像真正简化. 哥白尼发现将行星简单放置在日心天球上并不能满足观测的需要. 因此他像托勒密一样,不得不引进诸如本轮这样更为复杂的机制. 例如,哥白尼的计算显示出地球的(圆形) 轨道中心不是太阳,而是空间一点 C_E, C_E 沿着一个本轮旋转,而该本轮的中心 O 又以太阳为中心在旋转

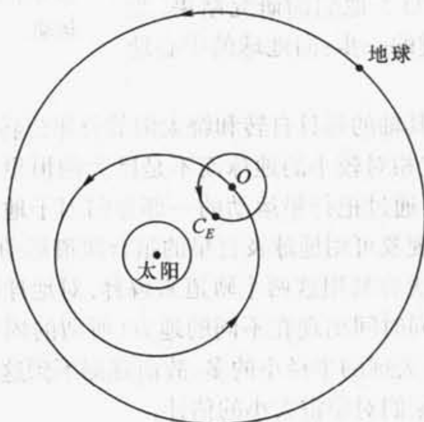


图 10.21　地球以 C_E 为心旋转, C_E 沿本轮以 O 为心旋转,而 O 以日为心在旋转.

(图 10.21). 类似地,其它各行星轨道的中心既不位于太阳处,甚至也不是地球轨道的中心. 结果,《天体运行论》所描述的整个体系与托勒密体系就复杂程度而言可以说是半斤八两.

哥白尼著作的数学细节使得大多数人无法问津,唯有当时最好的天文学家是其主要读者. 随后几十年中,那些数理天文学家发现应用哥白尼的理论和技术可以简化天文现象的计算,而对这些技术的使用并不必要去相信地球在运动. 因此,许多人(天文学家和受过教育的大众)仅仅将哥白尼的著作看成数学的假设,而不是物理学理论. 事实上,奥西安德(Andreas Osiander),这位路德教派的神学家在为《天体运行论》的印刷本所写的序言中就声称,哥白尼的地动说不应该被看成真的,而仅是计算上的假设,"因为那些真正的[天文] 法则是理性所不能及的."[33]



图 10.22　格列高里历法 400 周年纪念,梵蒂冈邮票.

然而,在 16 世纪后半叶,各种教会人士,特别是与罗马天主教会发生严重冲突的新教牧师,开始激烈反对哥白尼的思想,因为这些思想明显与《圣经》中断言地球稳定性的经文相抵触. 这些新教领袖相信罗马教会已远离《圣经》中所表达的观点. 他们感情冲动地排斥一切背离《圣经》文字的学说. 在此期间,天主教会本身倒几乎没有评论过哥白尼的著作. 事实

上,《天体运行论》在各种天主教大学中还作为教材被讲授,根据它而编制的天文表为历法改革提供了基础,该历法在1582年由主教格列高里十三世(Pope Gregory XIII)向天主教世界颁布(图10.22).具有讽刺意味的是,直到17世纪,当大多数天文学家通过新的证据及一种比哥白尼理论更好的日心说而相信地球确实在运动之后,天主教会却站出来竭尽全力反对地动说,并将其视为一切异端邪说的代表.

10.3.3 第 谷

有一位采用哥白尼著作作为天文计算基础的天文学家莱因霍尔德(Erasmus Reinhold,1511—1553).他在1551年公布了第一部三个世纪来在欧洲编制的天文表全集,这部天文表集通常因其保护人(普鲁士公爵)而被称之为"普鲁特尼克"(Prutenic)表.这些表明显地比过去的表好,部分原因是由于它们依靠了更多和更准的数据.然而,与基于托勒密著作的表相比其实并没有本质性的飞跃.在月食预测方面仍存在一天或更大的偏差.

不管怎样计算这些表,改进计算结果的关键是要进行更为精确的观测.第谷(Tycho Brahe,1546—1601)就是一位毕生献身于这种观测的天文学家.这当然需要先进的仪器及资金作基础.幸运的是,第谷于1576年说服丹麦国王腓特烈二世(Frederick II)允许他使用哥本哈根附近的汶岛(Hveen),并资助他建立了一所宏大的天文台,雇用必要的助手并采用新型仪器来进行常年观测(图10.23).第谷是最早认识到需要对各种行星进行连续观测的天文学家.虽然他最终离开了丹麦,移居布拉格(Prague)为奥地利皇帝鲁道夫二世(Rudolph II)工作,他还是积累了25年的大量数据,其精度已经达到几个弧分,远远超过以往任何最好的著作.



图10.23 第谷的天文台,含有一个象限仪及1572年的那颗新星,亚森欣岛的邮票.

其中有两项最重要的观测使他确信托勒密体系及作为其基础的亚里士多德哲学并不正确.首先,从1572年后期开始他对天空中出现的一个新的对象——一颗新星追踪了16个月.由于这一对象相对于星体的天球保持不动——第谷用精确的观测证实了这一点——他断定这一对象属于恒星的范围.因此,天空中的变化是完全可能的.这种可能性进一步被第谷于1577年观测到的一颗彗星所证实.通过这颗彗星的视差与月球和行星的视差的比较,他断言这颗彗星位于月球之外,并以太阳为中心、以大于金星离太阳的距离在旋转.因为在整个观测过程中,这颗彗星与太阳的距离有很明显的变化,第谷进而断言,天空不可能被镶嵌着行星的固体天球所充满.在行星之间必有真正的空间使其它天体也能够在其中运行.

10.3.4 开普勒和椭圆轨道

第谷主要是一位杰出的观测者,而非理论家.他曾设计了一个介于托勒密和哥白尼之间的折衷体系:即行星绕太阳运转,而太阳率领诸行星绕地球运转,地球仍是宇宙的静止不动的中心,但他未能对这一体系给出数学上的论证.开普勒(Johannes Kepler,1571—1630),在第谷生命的最后两年曾在布拉格为其工作,是一位有能力处理第谷的大量观测资料并构造出一种新日心说的天文学家,这种新日心说能够精确预测天文现象而不需要煞费苦心的本轮假设.

人物小传	开普勒(1571—1630)(Johannes Kepler)
	开普勒生于德国西南部的魏尔(Weil-der-Stadt),就学于蒂宾根(Tübingen)大学,在那里熟悉了哥白尼学说并确信它从根本上正确反映了世界体系.虽然他原想成为一名新教牧师,但命运发生了改变,蒂宾根大学推荐他去奥地利格拉茨(Graz)镇的一所学校填补一名数学教授的空缺.几年后当学校被关闭而新教职员都被流放时,开普勒却受到区别对待,允许他继续从事数学和天文学研究.开普勒明白要想完善哥白尼学说,他就必须了解第谷的观测结果.于是他开始与这位丹麦人联系,这导致他后来被鲁道夫二世皇帝任命为第谷在布拉格的助手.虽然在开普勒到达后大约 18 个月第谷就逝世了,但那时他已了解了大量第谷的工作,并能在自己的研究中使用这些资料了.他被任命为接替第谷的皇家数学家,在布拉格度过了随后的 11 年(图 10.24).



图 10.24　开普勒及其宇宙体系,匈牙利邮票.

也许是由于他的神学修养,加之哲学爱好,开普勒具有一种坚定不移的目标:寻求上帝用以创造宇宙的数学规则.正如他在 1596 年的最早著作《宇宙的秘密》(Mysterium Cosmographicum)开头所述,"量随物之始而生."[34] 在 1621 年第二版的一个注释中,他进一步阐明自己的意思:"量的概念永远与上帝并存,…….不信教的哲学家和教会的博学之士都同意这件事情."[35] 开普勒毕生致力于通过哲学的分析和繁复的计算来论证上帝创造宇宙的数量关系.他的目的似乎就是要在更高的层次上确认毕达哥拉斯的信条 —— 万物皆数.以哥白尼把太阳置于宇宙中心的学说为出发点,他最终发现了行星运行三大定律(即现今所称的开普勒定律),另外还有其它许多今天被我们视之为神秘主义而不予考虑的关系.

开普勒在《宇宙的秘密》中所详细讨论的这些关系之一是:为什么恰好有六个行星?因为"上帝是永恒的几何学家",这位至高无上的数学家想用正立体来分隔行星.欧几里得已经证明了仅有五种正多面体,于是开普勒便以此作为上帝恰好创造六个行星的原因.他然后设计出这样的想法,在每一对包含着相邻行星的轨道的天球之间有一个内接的正多面体(图 10.25).这样,在土星的天球内部有一个内接正六面体,而这个正六面体又外切着木星轨道.类似地,在木星和火星之间是正四面体,在火星和地球之间是正十二面体,在地球和金星之间是正二十面体,以及在金星和水星之间是一个正八面体.这些多面体位于天球之间的空间,且它们的大小为各行星轨道的大小关系提供了一种度量.例如,开普勒提出木星的轨道直径是火星的三倍,而正四面体的外切天球的直径恰好也是其内接天球直径的三倍.并非所有的数值都完全吻合,仍然存在偏差.但是这事实并没有对开普勒产生太大影响.他对这些数值为什么不如预期的那样准确给出了各种解释,其中一个理由是即使

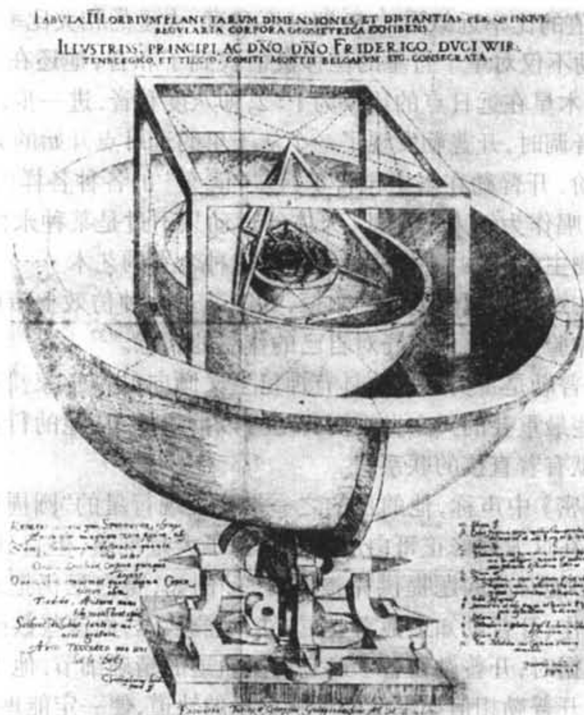


图 10.25　开普勒表示行星轨道的正则多面体，本图源于 1621 年版本的《宇宙的秘密》.（出处：Courtesy of the Department of Special Collections, Stanford University Libraries.）

普鲁特尼克表中的数据也并不完全精确. 他对自己的想法如此充满信心，那些偏差对他来说简直微不足道. 开普勒对这个问题的观点不能仅仅看成是他年青时的畅想. 事实上，他后来曾多次回到这个基本命题，每次都想为它的正确性引证新的理由.

开普勒对行星本身的大小比率也颇感兴趣. 他在《哥白尼天文学概要》(Epitome Astronomiae Copernicanae, 1618) 一书中指出，"[行星大小的] 数量级与天球的次序相同，没有什么比这与自然更和谐一致了."[36] 换句话说，水星应该是最小的行星而土星最大. 但是，这些大小之间的比率是什么? 他提出了几种可能性. 因为土星与太阳的距离近似于地球的 10 倍，他认为或者土星的直径是地球的 10 倍，或者它的表面积是地球的 10 倍，或者它的质量是地球的 10 倍. 为了作出选择，他参考了某些新的借助于望远镜获得的观测结果，并确认了第三种可能性. 由此可知，若土星的密度与地球的相同，则直径之比应该等于它们与太阳距离之比的立方根，而表面积之比应该是该立方根的平方. 开普勒没有证据能证明他的理论的正确性 —— 他没有办法测量行星的质量 —— 因此其论断仅仅是一种假说. 但是，就像讨论轨道大小的情形一样，开普勒确信必然存在一种简单的数量关系.

开普勒在音乐方面受过良好的教育，精通毕达哥拉斯的能定出和谐音调的弦长比率：1 : 2 是八度和音，2 : 3 是五度和音，3 : 4 是四度和音，等等. 在《宇宙和谐论》(Harmonices Mundi, 1619) 中，他试图把这些音律分配给与不同行星相关的各种数字. 首先，他尝试了公转周期，但没有发现任何和谐的比率. 其次，他尝试了行星的体积，最大和最小日距，速率的极值，行星在轨道上越过单位长度所需时间的变化. 结果是一无所获. 最后，在一段冗长的推演之后，开普勒总算找到了"正确的" 数字，

即从太阳处看到的行星每日视角位移. 土星的每日位移在远日点是 1′46″, 而它在近日点的每日位移是 2′15″. 因此, 这两个值的比率近似为 4 : 5, 即三度和音. 火星的相关比率是 26′14″ : 38′1″, 近似于 2 : 3, 即五度和音. 开普勒不仅对单个行星的位移极值找到了和音, 他还在不同行星的位移间发现了和音. 土星在近日点与木星在远日点的比率为 1 : 2, 即八度和音. 进一步, 当他将这样一组具体的比率转变为通常的键盘音调时, 开普勒发现了一个从土星的远日点开始的大调音阶, 和一个从土星的近日点开始的小调音阶. 开普勒在书中包含了行星"演奏"的各种各样的音符, 既有独奏又有合奏, 并以几段多部和声合唱作为结束: "因此, 天体的运动只不过是某种永恒的百音盒 ……. 这样, 就毫不奇怪: 人, 作为造物主的模仿者, 最终应该发现这种合奏的艺术 …… 以便能够借助由多种声音合成的艺术和音在某一短瞬表现出创造的永恒, 并能通过这种仿效上帝的音乐所发出的甜蜜欢乐的感觉在某种程度上体验上帝这位工匠对自己的作品的满意."[37]

人们有理由怀疑, 开普勒是否真是一位具有神秘主义倾向的科学家? 答案毫不含糊: 是. 然而, 开普勒就他所作出的一些最重要的天文发现而言是可以信赖的. 从他的行星运动三大定律到牛顿关于运动定律的基本贡献有着直接的联系.

开普勒在《宇宙的秘密》中声称, 他的目的之一是要发现行星的"圆周运动", 即确定它们的轨道. 他在这本著作中为了在这方面修正哥白尼体系进行了大量推演, 但直到该世纪末, 他终于认识到哥白尼的数学方案并不能对此问题提供完整的解答. 例如, 哥白尼仍把地球与其它行星区别对待. 为了修正哥白尼的著作, 开普勒知道他需要更好的观测数据, 而这些数据只能来自第谷. 到 1601 年, 当他终于拥有这些数据时, 开普勒便着手确定行星轨道的精确细节. 他先研究火星, 因为这颗行星的轨道一直最难把握. 开普勒相信如果他能理解火星的轨道, 便一定能理解所有行星.

在他 1609 年的《新天文学》(Astronomia Nova) 中, 开普勒向我们展现了他八年的详细计算、起点的失误和其它种种纰谬, 以及他对火星轨道计算的不懈努力. 他首先决心要算出地球轨道的精确参数, 因为他觉得自己胜哥白尼一筹之处, 正是在于认识到火星的运行是从转动的地球上观测的. 开普勒假设地球的轨道是中心在 C 点、半径为 r 的圆, 而太阳在 S 点且 $CS = e = 0.018r$ (图 10.26). (地球的轨道非常接近于圆形, 故假设它是圆并未把开普勒引入歧途.) 此外, 在直径 CS 上另有一点 A, $AC = CS$, 且 $\angle EAQ$ 随时间而匀速变化, 其中 Q 是地球的远日点. 换句话说, 他重新引入了被哥白尼所放弃的等距点. 因为地球在轨道上关于 A 点以固定不变的角速率运行, 所以它的线速率必会随其日距的变化而变化. 开普勒发现地球在远(近)日点附近的线速率与日距成反比, 然后便将其推广到轨道的其余部分. (不幸地, 这个法则后来被证明并不正确. 正如他后来所知道的那样, 只是行星速率在垂直于半径向量上的分量与它的日距成反比.)

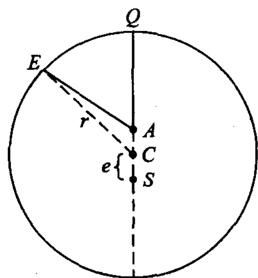


图 10.26　开普勒关于地球轨道的假设.

然而, 不同于托勒密及哥白尼, 开普勒不仅关心天体运行的纯数学计算(就是说不只关心"治表"), 而且还关心其物理解释. 他试图描绘地球在空间的实际轨道, 故而想知道是什么原因引起地动, 又是什么将它维系在轨道上, 以及为什么速率会随日距而变化. 在读了吉尔伯特(William Gilbert) 的《磁体论》(On Magnets, 1601) 之后, 开普勒确定了这样的事实, 即由太阳发出的某种力作用于行星之上, 并驱使它沿其轨道运行. 他觉得这种作用在地球上的力使地球沿圆周运动比作用在行星上的力使其沿本轮运动更容易让人理解. 另外, 像磁力一样, 太阳的力随距离越远而变得越弱, 故而行星的速率在远处而变小. 这种从数学观点向物理观点的改变, 正是开普勒在重新引入等距点而放弃本轮说时感到惬意自得的原因之一.

回到火星的运动，开普勒起初还使用他早期的假设，即火星轨道的圆周性，因为这个假设至少可以提供比较近似的结果．他的目的是要计算火星在达远日点 Q 之后所运行的弧长 QP 与经过这段弧所用的时间之间的关系．他知道火星距日越远运行得越慢．然而，由于速率与弧长之间的精确关系的计算是他力所不及的，开普勒只好借助于近似．他关于行星的速率与其向径的长度成反比的假设，对地球及火星都适用，这个假设意味着通过(无穷小的)弧段所需的时间与其向径成正比．那么，只要选取适当的单位，时间便可用向径来表示．开普勒然后推理道，通过有限弧段 QP 所需的整个时间可以被看作构成那部分圆的向径之和，或即向径所扫过的面积(图 10.27)．开普勒知道这样的无穷小推理并不严密，但他还是把它陈述为一个法则(一个基于不正确的圆形轨道和不正确的速率法则而得出的法则)：行星的向径在相等的时间内扫过的面积相等．这个法则通常称之为开普勒第二定律，因为我们今天把它看成是第一定律的补充．非常有趣的是，当开普勒后来发现正确的行星轨道不是圆而是椭圆之后，他并未试图重新证明这一定律．

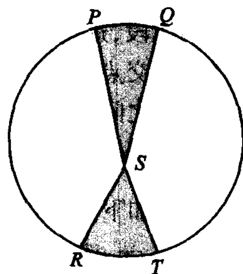


图 10.27 开普勒第二定律：行星的向径在相等的时间内扫过的面积相等．当面积 SPQ 等于面积 SRT 时，行星从 Q 到 P 的运行时间等于从 R 到 T 的运行时间．

行星运行轨道的椭圆性是开普勒第一定律的内容．开普勒也详细说明了他是怎样发现这一定律的．在算出地球轨道之后，他对火星到所假设的轨道中心的距离进行了各种计算，并发现算出的距离在远日点和近日点附近较大而在轨道的其余部分较小．因此，轨道不可能是圆周．开普勒断定轨道只能是某种卵形线．要放弃希腊人的令人赏心悦目的圆形线并代之以模糊不清的卵形线，这多少会让人感到有些不可思议，因为这种曲线会从外观上摧毁开普勒一直在探索的"天球和谐"的一切可能性．然而，开普勒还是开始了计算这种卵形线的精确形状的漫长过程．

经过两年的计算，情况渐趋明朗．为了帮助计算，他一直在用椭圆接近卵形线．他注意到圆周与椭圆短轴终点之间的距离 AR 等于 0.004 29 (圆的半径设为 1)，这个值其实是 $(1/2)e^2$，其中 $e = CS$，是圆心与太阳的距离(图 10.28)．由此可知，

$$CA : CR = 1 : 1 - \frac{e^2}{2} \approx 1 + \frac{e^2}{2} = 1.00429.$$

这个数字似曾相识，这引起了开普勒的深思．它等于 5°18′ 的正割值，即方向 AC 和 AS 的夹角 ϕ 的正割值，其中 A 是与远日点 Q 成 90° 的圆周上的点．这个正割值在目前情形就是向径的长度与其在一

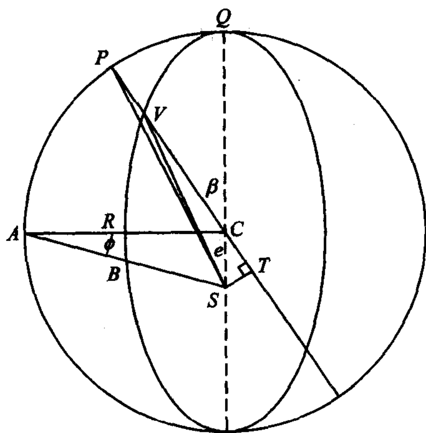


图 10.28 开普勒椭圆轨道的由来．

直径上投影的比率.认识到 $CA : CR \approx SA : SB \approx SA : CA$ 之后,开普勒便产生了一个闪光的灵感,当 CQ 与方向 CP 之间的夹角为任意角 β 时(不一定正好 90°),距离 SP 与太阳—火星的实际距离之比也是 SP 与它在一直径上垂直投影 PT 之比.换句话说,他认识到了太阳—火星的实际距离是 PT,而 $PT = PC + CT = 1 + e\cos\beta$.对开普勒来说剩下的问题是怎样标出这个距离.开普勒最初决定用一端在太阳处而另一端在向径 PC 上的线段来表示它,即让 $SV = PT$.不幸的是,如此描绘的曲线最终与观测不符,事实上,它并不是一个真正的椭圆,而开普勒此时由于各种原因已认定真正的轨道应该是椭圆.

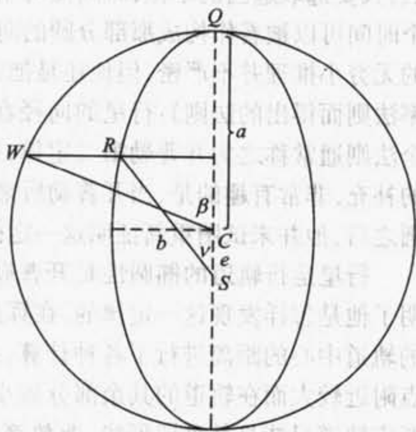


图 10.29 开普勒证明轨道曲线是椭圆.

开普勒最终发现了正确的结果,即由 $\rho = 1 + e\cos\beta$ 给出的距离应该表示成起点为太阳、终点在 CQ 的一条垂线上的线段,这里 β 是 CQ 和 CW 的夹角,而 W 是该垂线与辅助圆的交点(图 10.29).(应该注意,此结果与开普勒最初的想法之间只有极小的差异,偏差不超过 5 弧分.)开普勒已能证明他现在所建立的曲线是椭圆,其推理可用现代记号概述如下.假定椭圆中心为 C,且 $a = 1, b = 1 - \frac{e^2}{2}$,其中 $e = CS$.这个椭圆可以认为是由半径为 1 的圆对所有垂直于 QC 的纵坐标按照比例 b 经压缩而成.如果 ν 表示弧 RQ 在点 S 的对角,那么 $\rho\cos\nu = e + \cos\beta$ 且 $\rho\sin\nu = b\sin\beta$.将这两个方程的两边平方并相加,得

$$\rho^2 = e^2 + 2e\cos\beta + \cos^2\beta + \left(1 - \frac{e^2}{2}\right)^2\sin\beta$$
$$= e^2 + 2e\cos\beta + 1 - e^2\sin^2\beta + \frac{e^4}{4}\sin^2\beta.$$

忽略不计带 e^4 的项,则有结果

$$\rho^2 = 1 + 2e\cos\beta + e^2\cos^2\beta = (1 + e\cos\beta)^2.$$

椭圆方程可表示为 $\rho = 1 + e\cos\beta$,这与已经导出的轨道本身的曲线方程完全一样.另外,如果再次忽略不计带 e^4 的项,则椭圆中心到焦点的距离 c 为

$$c^2 = 1 - b^2 = 1 - \left(1 - \frac{e^2}{2}\right)^2 = e^2,$$

由此便知太阳位于椭圆的一个焦点且 e 是离心率(图 10.30).我们现在已导出了开普勒的行星运动第一定律:行星绕太阳沿椭圆运行,太阳位于椭圆的一个焦点上.开普勒本人对火星导出这个定律后,仅仅通过对其它行星的简单检验便断定了它的一般有效性.[38]



图 10.30 开普勒作出的椭圆轨道,德国邮票.

开普勒的第三定律作为一个经验事实首次出现在《宇宙和谐论》中.在某种意义上说,它是早在《宇宙的秘密》中已经开始的一系列研究的顶点,因为它为开普勒所提出的关于轨道大小和运行的一般性问题提供了另一个答案:"可以完全肯定和确切地说,任意两个行星公转周期的比率等于行星日距平均值的 3/2 次幂之比."[39] 开普勒对第谷的测量做了更多的研究之后发现了这个定律,但从未用其它原理来给出它的证明.

开普勒的行星运行三大定律在天文学与物理学理论的发展上具有巨大的影响.它们的发现对科学家使用的方法论也提供了一个杰出的范例.科学家开

始需要一些理论,但随后必须始终对理论结果和观测结果进行比较.如果他们对观测有信心,而后者又与理论预报不相符,他们就必须修改理论.开普勒经常这样做,直到他最终达到理论结果与观测结果的一致.他花费了好几年时间来完成必要的计算.然而,在他临去世时,对数的发明却大大地简化了他和其他天文学家的计算.

10.4 对 数

对数思想可能来源于某些将乘法化为加减法的三角公式.我们知道用正弦定理解三角形时,需要进行乘除运算.由于正弦通常要算到七位或八位数字(使用半径为 10 000 000 或 100 000 000 的圆),这些计算冗长并经常出错.天文学家认识到若用加减法替代乘除法,则可方便计算并减少错误.为此,16 世纪的天文学家经常使用诸如 $2\sin\alpha\sin\beta = \cos(\alpha-\beta) - \cos(\alpha+\beta)$ 之类的公式.例如,要用 4 378 218 乘以 $\sin\beta$,其中 $\beta = 27°15'22''$,可以先确定 α,使之满足 $\sin\alpha = 2\,189\,109$,然后用表确定 $\cos(\alpha-\beta)$ 和 $\cos(\alpha+\beta)$.后两值的差即为所求的乘积,其计算过程未含任何乘法.

另外,更明显地,对数思想的根源也可在斯蒂弗尔和许凯这样的代数学家的著作中找到,他们的书中都列有 2 的幂与其指数的对照表,并说明了其中一个表的乘法对应另一个表中的加法.但由于这些表值的间隔越来越大,它们并不能被用于所必要的计算.然而,在 17 世纪转折时期,苏格兰的纳皮尔(John Napier,1550—1617)和瑞士的比尔吉(Jobst Bürgi,1552—1632)各自独立地提出了编制一种推广数表的思想,这种表能通过加法来完成任意两数(不仅是 2 的整数幂)的乘法运算.纳皮尔首先出版了他的著作.

10.4.1 对数的思想

纳皮尔的对数表最早出现在 1614 年的著作《奇妙的对数法则的说明》(Mirifici Logarithmorum Canonis Descriptio)中.这本著作仅包含了一个如何使用对数表的简要介绍.他的第二本关于对数的著作《奇妙的对数法则的构造》(Mirifici Logarithmorum Canonis Constructio),在他逝世两年之后于 1619 年出版,其中陈述了造表所依据的理论.这后一著作阐明了他的富于想象力的思想,即用几何学构造一个表来改进算术.

考虑到天文学家的计算主要涉及三角函数,特别是正弦,纳皮尔便致力于构造一个表把正弦的乘法转化为加法.为了定义对数,纳皮尔构想出两条数字线.在一条线上标有一个递增的算术序列,$0, b, 2b, 3b, \cdots$,而在另一条线上标有一个数列,这个数列与右端点的距离形成一个递减的几何序列,$ar, a^2r, a^3r, \cdots$,其中 r 是第二条线的长度(图 10.31).(纳皮尔选择 r 为 10 000 000,即他的正弦表中圆的半径,a 为小于 1 但非常接近于 1 的一个数(补遗 10.1).)这条线上的点可以记为 $0, r-ar, r-a^2r, r-a^3r, \cdots$.对于纳皮尔来说,这些点一般表示某些角的正弦值.

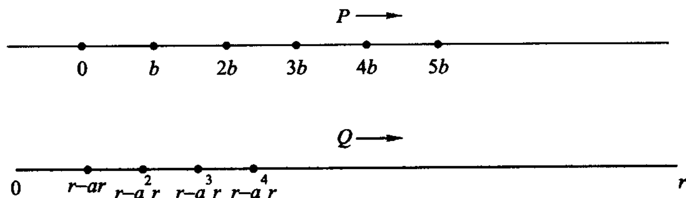


图 10.31 纳皮尔的移动点.

补遗 10.1

近代十进小数记号

纳皮尔对近代十进小数记号有重要贡献.斯蒂文已经详述过这种思想,并提出一种记法.但是,纳皮尔在《构造》一书的开始部分指出,精确的计算需要使用像 10 000 000 这样的大数作为正弦表的基数,他接着写道:"在造表中,在这样的大数后添加一个句点和若干阿拉伯数字就可以表示一个更大的数.…… 这样用句点分隔的数,不论怎么写,句点之后总表示一个分数,其分母是单位 1 后再加若干个零,而零的个数与句点之后的数字位数一样多."[40] 例如他写道:25.803 与 $25\frac{803}{1000}$ 相同,9 999 998.000 502 1 是指 $9\,999\,998\frac{5021}{10\,000\,000}$.含有这些十进小数的纳皮尔正弦表的出版,不久便导致它们在整个欧洲的广泛使用.自从印度 - 阿拉伯数字传入欧洲以来,完备的十进制位值系统获得普遍采用大约经过了 400 年之久.

纳皮尔现在来分析点 P 和 Q 在每条线上向右的移动:P 在上面的线上"算术地"移动(亦即,具有固定不变的速率).因此,P 在相同的时间内,通过各个相等的间隔 $[0,b]$,$[b,2b]$,$[2b,3b]$,….Q 在下面的线上"几何地"移动.它的速率是变化的,以便在相同的时间内,通过各个递减的间隔 $[0,r-ar]$,$[r-ar,r-a^2r]$,$[r-a^2r,r-a^3r]$,….在每一间隔的位移形成一个递减的几何序列 $r(1-a)$,$ar(1-a)$,$a^2r(1-a)$,… 这个序列中每一项都是相应的间隔的左端点到整个线段的右端点的距离的相同倍数.因为相等时间内的位移与速率一样具有相同的比率,由此便知该点通过每一间隔的速率与该间隔的开始到整个线段的右端点的距离成比例.纳皮尔起初似乎认为下面线上的点的速率在它经过每一标点时是突然变化的,而在每一给定间隔的其余部分保持不变.然而,在他的对数定义中,纳皮尔通过把第二个点的速率视为连续变化而消除了这种突变(当然,他并未使用这种术语).因此,如果一个点的速率与该点到整个线段右端点的距离总是成比例,则该点为几何移动.纳皮尔说,"某一给定的正弦值的对数是一个始终以相同的速率算术增加的数,而半径是以该速率开始几何减少的,并在相同的时间内半径减少到所给定的数."[41] 换句话说,如果上面的点 P 和下面的点 Q 都从 0 点开始以同样的初速移动,点 P 是匀速移动,而点 Q 是几何移动,再假设当 Q 到达一个点,该点到右端点(即半径)的距离为 x 时,点 P 到达 y,那么定义 y 是 x 的对数.

按照现代微积分学记号,纳皮尔的思想反映在如下的微分方程中

$$\frac{dx}{dt} = -x, x(0) = r; \quad \frac{dy}{dt} = r, y(0) = 0.$$

第一个方程的解是 $\ln x = -t + \ln r$,或 $t = \ln\frac{r}{x}$.比较这个解与第二个方程的解 $y = rt$,说明纳皮尔的对数 y(此处记为 $y = \text{Nlog } x$)可表示成现代自然对数的形式 $y = \text{Nlog } x = r\ln\frac{r}{x}$.因此,纳皮尔的对数密切相关于自然对数.然而,它并不具有普通自然对数的性质,例如,它的值随着 x 值的增加而减少.

10.4.2　对数的应用

虽然纳皮尔的定义有些不同于现代形式,但他还是导出了一些类似于现代对数的重要性质,并说明了怎样构造正弦的对数表.他开始便指出,由定义可知 $\text{Nlog } r = 0$,因为上线的点完全未动.事实上,纳皮尔知道可以把任一固定数的对数指定为 0,但是他说,"最好取其为整个正弦值[即圆的半径],正弦的对数的加减法在各种计算中最常见,这样对我们更方便."[42] 类似地,若 $\alpha/\beta = \gamma/\delta$,则 $\text{Nlog } \alpha - \text{Nlog } \beta = \text{Nlog } \gamma - \text{Nlog } \delta$.这个结论也可从定义得出,因为下线的点的几何移动就意味

着从 α 到β的时间等于从γ到δ的时间.由此可推出一些在计算中应用对数的法则.例如,若 $x:y=y:z$,则 $2\text{Nlog}\, y=\text{Nlog}\, x+\text{Nlog}\, z$;若 $x:y=z:w$,则 $\text{Nlog}\, x+\text{Nlog}\, w=\text{Nlog}\, y+\text{Nlog}\, z$.另一方面,纳皮尔没有说明如何计算乘积的对数,也许他对这样的纯乘法运算不感兴趣.他是针对三角学而建构对数的,解三角形的许多运算需要求比例第四项,他的法则对此确实有用.

作为这种计算的一个例子,设直角三角形的斜边 c 和一直角边 a 为已知,求 a 的对角 α.纳皮尔使用基本三角关系

$$\frac{\sin\alpha}{r}=\frac{a}{c},$$

其中 $r=10^7$ 是定义正弦所用的圆的半径.然后,应用他的表以及上述比例法则来计算

$$\text{Nlog}\sin\alpha=\text{Nlog}\, a-\text{Nlog}\, c+\text{Nlog}\, r.$$

由于 $\text{Nlog}\, r=0$,α 的正弦的对数便表示为相应边的对数.反查他的对数表即得所求的角.虽然纳皮尔的表是正弦的对数表,但他在这个问题中也用它来求所需边长的对数,做法是对该数的数字作适当调整,并在表中找一个最接近于该数的正弦值,然后取该正弦值的对数.

纳皮尔给出了应用该表的许多其它例子.为解已知两边 a、b 及另一边 a 的对角 α 的平面三角形,纳皮尔将他的对数法则应用于正弦定律

$$\frac{\sin\alpha}{a}=\frac{\sin\beta}{b},$$

注意,其中 β 有两个可能的值,一个锐角和一个钝角.为解已知两边 a、b 及其夹角γ的三角形,纳皮尔没有使用作垂线的标准方法,因其不适合于对数计算.取而代之,他采用了正切定律:

$$\frac{a+b}{a-b}=\frac{\tan\frac{1}{2}(\alpha+\beta)}{\tan\frac{1}{2}(\alpha-\beta)}.$$

若已知 γ,则可求出 $\alpha+\beta$.对这个比例取对数可求出 $\tan\frac{1}{2}(\alpha-\beta)$,进而$\frac{1}{2}(\alpha-\beta)$,以及 α 和β均可求出.纳皮尔是怎样用他的正弦的对数表计算正切对数的?为了回答这个问题,我们来列出一行纳皮尔的表,纳皮尔表对从 0° 到 45° 每隔一弧分的角给出了七列数字:

34°40′　5 688 011　5 642 242　3 687 872　1 954 370　8 224 751　55°20′.

第一列给出弧(或角)的值,而第二列给出该弧的正弦.最后一列给出的是与第一列弧互余的弧,而第六列给出它的正弦.由此便知,第六列给出了第一列弧的余弦值.第三和第五列依次给出第二和第六列的正弦的纳皮尔对数值.最后,中间一列表示第三和第五列表值的差,或第一列弧的正切的纳皮尔对数.因为 10 000 000 的对数是 0,大于 10 000 000 的数的对数就必须是负的,其定义可以通过在原始定义中简单地反转移动点的方向来给出.当然,这些数不能表示正弦,而是表示正切或者正割.在这种情况下,中间一列的对数的负数是 55°20′ 的正切的对数,而第三列的对数的负数是同一角的正割的对数.

虽然仅从纳皮尔的运动定义中我们不能详述他的对数表的实际建构过程,但需要指出的是这个过程花费了他20年.[43] 即使这项成果诞生于手工计算时代,其中几乎没有错误存在.然而,在他晚年,纳皮尔断言如果把 1 而非 10 000 000 的对数取为 0,这样会更方便.在这种情况下,熟知的对数性质 $\log xy=\log x+\log y$,$\log x/y=\log x-\log y$ 将成立.进一步,如果 10 的对数取为 1,则 $a\times 10^n$ 的对数,其中 $1\leqslant a<10$,就简单地等于 a 的对数加上n.纳皮尔在能够建构出一个以这些原则为基础的新表之前就去世了,但是,曾在 1615 年与纳皮尔详尽讨论过个问题的布里格斯(Henry Briggs,1561—1631),却开始了这种表的计算.不过,布里格斯并非简单地用算术过程把纳皮尔的对数转换

成"常用"对数,而是从头开始制作该表.既然 $\log 10 = 1$,他相继计算了 $\sqrt{10}, \sqrt{\sqrt{10}}, \sqrt{\sqrt{\sqrt{10}}}, \cdots$,直到 54 重开平方根之后,他得到一个非常接近于 1 的数.所有这些计算都进行到 30 位十进数.由于

$$\log\sqrt{10} = 0.5000, \log\sqrt{\sqrt{10}} = 0.2500, \cdots, \log(10^{1/2^{54}}) = 1/2^{54},$$

这使他能够通过对数法则造出一张分布较密的对数表.布里格斯的表于 1628 年由费拉克(Adrian Vlacq)完成,成为随后几乎一切对数表的基础.天文学家很快发现了用对数进行计算的巨大优点.对数变得如此重要,以致 18 世纪的法国数学家拉普拉斯(Pierre-Simon de Laplace)声称对数的发明,"以其节省劳力而使天文学家的寿命延长了一倍."

10.5　运动学

我们将要讨论的迪伊的最后一种数学艺术涉及运动."静力学是一门数学艺术,它揭示万物轻重之原因,探究轻重事物的运动和性质,"[44] 而"圆动学(Trochilike)…… 探究各种简单的和复杂的圆周运动的性质."[45] 伽利略(Galileo Galilei, 1564—1642),这位近代物理学的奠基者,在重建古希腊人

人物小传

伽利略(1564—1642)(Galileo Galilei)

伽利略(图 10.32)于 1581 至 1585 年就学于比萨大学,形式上攻读的是医学.然而,他对数学更感兴趣,最后没有获得任何学位便离开了大学.他在数学方面接受的是古典教育.他熟知亚里士多德和欧几里得,并读过一些阿基米德的著作.因此,他精通欧多克索斯的比例理论,却对卡尔达诺和邦贝里的代数及更近的韦达的工作知之甚少.他深信在自然现象的研究中数学、特别是几何学的重要性.

如今伽利略最出名的是他与天主教会之间围绕其著作《关于两大世界体系的对话》(Dialogue Concerning the Two Chief World Systems, 1632)所发生的冲突.这部著作阐述了他对托勒密和哥白尼宇宙论的正反两方面的分析.正如本章开始所述,教会当局已在 1616 年警告过伽利略,对地球的静止不动不能持有相反的观点,并声明这是教会官方的立场.因此,伽利略很谨慎地把哥白尼的观点作为假设来表述,并同时介绍了哥白尼观点的相应推论和传统的托勒密观点的推论.尽管如此,仔细研读此书便会发现,其实伽利略深信地球是围绕太阳旋转的——这在当时已并非令人吃惊的结论,而他的写法使旧观点的拥护者们显得愚蠢可笑.伽利略仍旧相信科学与宗教真理是能够调和的,正如他在 1615 年所写的那样:"当我们在物理学中得到任何确定的事实时,我们就应当利用它们作为最佳帮助来正确诠释《圣经》,并研究蕴涵其中的真正含义,这些与已经证明的真理之间应该是一致的."[50] 不幸,1630 年代的宗教领袖们却冥顽不化,坚信对《圣经》的流行解释的任何挑战都应被坚决抵制.于是伽利略在 1633 年被送上了罗马宗教法庭,并被迫忏悔认错.其后,他被判终身软禁并禁止再出版任何著作.然而,他后来还是设法将他的另一本最重要的著作《关于两种新科学的对话与数学证明》(Discourses and Mathematical Demonstrations Concerning Two New Sciences)的手稿送到处于宗教法庭管辖范围之外的荷兰莱登,在那里由埃塞菲尔斯(Elseviers)出版社于 1638 年出版.

虽然教会禁止《对话》,但该书当时已经广为流传,影响难以消除.因此,意大利学界以及其它地区的读者,不久便相信哥白尼体系中的地动说确属真理.最终,甚至教会也不得不承认他们对《圣经》某些内容的解释应该改变.

及后来某些中世纪学者曾探讨过的运动定律方面做出了很大贡献.但是像其前辈们一样,他是用几何而不是代数的语言来阐释自己的思想的.虽然他的研究工作属于我们今天一般所称的静力学范围,但他最重要的新思想则涉及自由落体的"自然"加速运动和抛射体的"弯曲"运动,并发表于他1638年的著作《关于两种新科学的对话与数学证明》中.因此,正如哥白尼和开普勒应用数学研究天体运动一样,伽利略则是应用数学来研究地球上的运动.



图 10.32 意大利邮票中的伽利略.

10.5.1 加速运动

《两种新科学》是以三人对话的形式写成,对话围绕着一篇欧几里得式的运动学论文框架而展开,该论文包含一个定义,一个公设,若干定理和证明.定义说,"均匀加速的运动是指在相等的时间内其速度的增量相等."[46] 尽管伽利略是从与他的中世纪前辈相同的定义出发,但他取得了两点重大的进步.其一,到1604年他已发现自由落体是匀加速运动;其二,他得出了这个事实的一系列数学推论,其中有些可用实验证实.[47]

伽利略曾相信自由落体的速度与下落的距离而不是与所用的时间成正比.在《两种新科学》中,他给出一个论证来说明前者是错误的.首先,他指出:如果一给定物体的二个不同速度与该物体依次取此二速度时分别通过的距离成正比,则该物体通过相应距离所用的时间相等.这一断言对于在给定的时间段内速度为常数的情形是显然成立的.伽利略然后假定它对速度连续变化的情形亦成立.于是,"如果自由落体下落4布拉西亚(braccia,意大利长度单位)时的速度是下落2布拉西亚时的速度的两倍的话,由于一个下落距离是另一个的两倍,故通过这两段距离所用时间应该相等."[48] 伽利略在这里比较了两个(无限的)速度集,即落体通过(离原点)头一个2布拉西亚路程内每一点的瞬时速度的集合与它通过头一个4布拉西亚路程内每一点的瞬时速度的集合.断言说通过这两段距离的总时间相等,相当于要将此结论应用于无限多个无限小的时间区间,并对所有这些无限小时间区间求和.伽利略总结道,初速为零的自由落体不可能在相同的时间内,既下落2布拉西亚又下落4布拉西亚,故速度与距离成正比是错误的.

伽利略这种比较两个无限小量集合的论证方法在数学史上是破天荒第一次,但他在其它场合特别是在证明中世纪的平均速度法则时也使用过这种方法.

定理 初速度为零的匀加速运动物体通过给定距离所用的时间等于该物体以某一匀速运动通过相同距离所用的时间,此匀速运动的速度是前述匀加速运动的最大速度[即最后一瞬的速度]的一半.[49]

用 AB 表示移动时间,EB 表示运动物体达到的最大速度,F 为 BE 的中点,伽利略构造出直角三角形 ABE 和矩形 $ABFG$,它们的面积相等(图10.33).因此,一方面在 AB 的点所表示的时刻与三角形中表示速度增加程度的平行线之间,另一方面也在那些时刻与矩形中表示匀速(等于最终速度之一半)的平行线之间,存在着一一对应关系.伽利略总结道,由于中点上方的亏缺三角形与下方的盈余三角形相补,"相应于三角形 AEB 中不断增大的平行线的加速运动的速度瞬量与相应于矩形 $ABFG$ 中的平行线的等速运动的速度瞬量一样多."[51] 因为在每一时刻的速度瞬量与在此时刻内通过的距离成正比,故两种情况下的总位移相等.

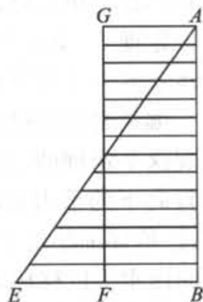


图 10.33 伽利略对平均速度定理的证明.

作为这个定理的推论,伽利略证明了对于初速为零的自由落体,距离与

时间的平方成正比.亦即,$d_1 : d_2 = t_1^2 : t_2^2$.用现代术语可表示为 $d = kt^2$,但伽利略用的是欧几里得的比例概念,而非现代的"函数"概念.为了证明此推论,伽利略首先指出对于两个不同的匀速运动而言,$d_1 : d_2 = (v_1 : v_2)(t_1 : t_2)$.根据平均速度定理,自由落体在两个时间 t_1 和 t_2 内的位移 d_1 和 d_2 分别等于其最大速度之半的匀速运动在相同时间内的位移.而最大速度之半也与时间成正比,故上述复合比率中的速度之比可替换为时间之比,推论得证.

伽利略大约证明了关于匀加速运动的 38 个命题.除了自由落体外,他也对斜面运动中速度、时间和距离的关系很感兴趣.通过实验证实小球沿斜面无摩擦滚动时,速度和时间成正比,距离和时间平方成正比,这就是匀加速运动的定律.当小球沿一个光滑斜面滚下之后,可以沿着另一个斜面上升到小球原来的高度,而与斜面的斜度无关,如果第二个斜面是水平面,小球将沿着这个平面一直运动下去,这就是惯性原理.伽利略对于求解最速降线问题也有所研究.即寻求一条路径,使物体在最短时间内从一个点下降到另一个低处的点.他演示了在一个垂直的圆上,物体沿着一条弦 DC 下滑所用的时间要比沿着两条弦 DB, BC 下滑所用的时间长(图 10.34).(此处 DC 必须对应一个不大于 90° 的弧线.)通过无穷小分析,即把这个结果扩展到越来越多的弦,他得出一个错误的结论,最速降线是圆弧.直到 17 世纪末才有几位数学家推演出这个曲线事实上应是旋轮线.

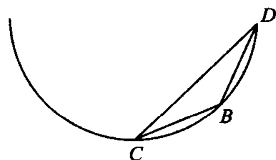


图 10.34　伽利略和最速降线问题.

10.5.2　抛射体运动

在《两种新科学》的最后部分,伽利略研究了抛射体的运动.他发现,抛射体运动可以分解为两个部分,一个是沿着水平方向的匀速运动,另一个是垂直方向的匀加速运动.他写道:

> 我设想把一个小球投掷到一个水平面上,如果所有阻力忽略不计,且平面是无限延伸的,…… 则这种匀速运动状态将永远保持不变;但如果平面有界且被置于高处,则当小球脱离该平面之后,由于永远向前的匀速运动和因小球自身重量而产生的垂直向下的匀加速运动的共同作用,便复合产生一种新的运动,我故且称其为抛射体运动.[52]

伽利略此处陈述了惯性定律,即小球在无摩擦的水平面上运动时,其匀速状态永远保持不变,正如他以前所述,因为"没有加速或减速的原因."[53] 牛顿把伽利略的"原因"替换为他自己的外力作用后,便将这一原理推广为他的力学基本定律之一.然而,伽利略的兴趣却不在定律本身,而是抛射体的运动路径.他证明了

定理　由水平方向的匀速运动和垂直向下的匀加速运动复合生成的抛射体运动,其路径是一条半抛物线.

伽利略是通过他 1608 年所做的一个桌面滚球实验而发现这个定理的,该实验使他相信水平方向的运动不受垂直向下的重力运动的影响.[54] 他在证明中应用了这个假设.他详细绘出了一个抛射体的路径图,指出在相等的时间内水平位移相等,而在相等的时间内,垂直位移与时间的平方成正比.因此,这个曲线具有以下性质,对于其上的任意两点 F、H,水平位移的平方比 $FG^2 : HL^2$ 等于垂直位移的比 $BG : BL$(图 10.35).由于伽利略对阿波罗尼

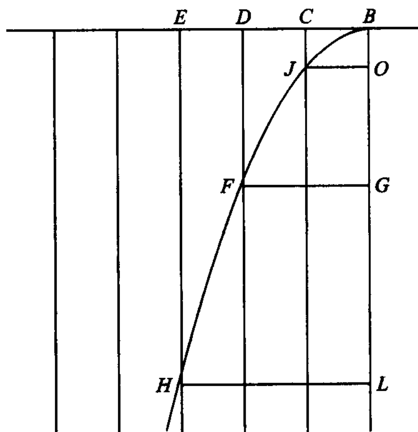


图 10.35　伽利略和抛射体的抛物线运动.

奥斯的著作很熟悉,他便直接得出这条曲线是抛物线.

伽利略还进一步证明了以任何角度发射的炮弹,其路径也为抛物线.事实上,针对不同的发射仰角,他计算了若干炮弹的高度和距离,还发现初始角为45°时,能达到最大的射程.塔塔利亚在他1537年的著作《新科学》中已给出了后一结论,但并未指出路径是抛物线.他甚至还在伽利略之前指出了,以互余角发射的炮弹将具有相等的射程.

但与伽利略不同的是,塔塔利亚的研究并非基于任何物理学的基本原理.他也没有像伽利略那样真正理解数学模型的作用.而伽利略写道,“还没有哪门严格的科学能用来研究重量、速度及变化莫测的状态.因此要科学地处理这样一些问题,就有必要对它们进行抽象.我们还必须对在有阻力的情况下抽象出来的结论进行论证和修正,以使它们也能适用于经验所遇到的各种实际情况.……确实,实际的抛射体路径与精确的抛物线之间的偏差就几乎难以察觉.”[55] 因此,伽利略表达了他对数学应用物理学的坚定信念.人们必须通过对给定情形中最重要的概念进行分析,以形成数学模型.仅当在数学上推导出该模型的一系列结论并经过与实验结果的比较,才能决定是否需要对模型进一步修正.像开普勒一样,伽利略始终遵循这种将物理现象数学模型化的基本方法.因为开普勒研究的是天文现象,他只能将理论结果与观察作比较.而伽利略却是通过实验来证实(或否定)其理论结果.对这种数学模型化过程的详细阐释甚至比实际的物理定理本身还重要,这正是伽利略在数学和物理互相发展方面最基本的贡献.当17世纪的科学革命在牛顿的著作中达到最高峰时,伽利略的这些思想终于开花结果,大放异彩.

习　题

透视问题

1. 用透视图法画出一个方格图案.首先用适当的距离确定出没影直线和没影点,然后用课本中的法则建构出水平线.
2. 阅读 Julian Lowell Coolidge 论述弗兰切斯卡的著作 *The Mathematics of Great Amateurs* 中的第三章,并简述弗兰切斯卡关于空间立方体的画法.
3. 收集一些文艺复兴时期的油画,包括弗兰切斯卡,阿尔贝蒂和迪勒.标出其中的没影点和没影线,并指出一些汇聚在没影点处的直线.

地理和航海问题

4. 写出一个关于投影地图的简要报告.对目前最常用的几种投影进行数学描述,包括球极投影,锥顶投影和柱面投影.
5. 本题给出建构墨卡托投影图的实际过程,来表示赤道与北纬30°以及西经75° - 85°之间的区域.用10 cm的直线表示上面两子午线之间的赤道.将它分为1 cm长的小区间,并准备过分点画垂直于赤道线的子午线,那么1°的经线长为1 cm.为了求出北纬10°平行线在地图上的距离,由于1 cm对应于1°的大圆,相应的球体半径必须是$180/\pi$,因此应该用这个值乘以$D(10°)$,而$D(10°)$的计算见公式10.1.类似地,为计算20°平行线与10°平行线之间的距离,应该求出$D(20°) - D(10°)$,再乘以该半径.还要算出30°平行线的距离.为了制图更精确,可再算出5°,15°和25°平行线的距离.
6. 修改习题5关于平行线的计算,以便绘出西经80° - 100°以及北纬40° - 60°之间的区域,假设北纬40°平行线上1 cm长对应于1°的经度.

三角学问题

7. 求解课本中所述《论各种三角形》中的问题：已知三角形 ABG 的底边 $BG = 20$，垂线 $AD = 5$，比率 $AB : AG = 3 : 5$，求两边 AB 和 AG（见图10.15）.
8. 本题及下一道题也出自《论各种三角形》. 在三角形 ABC 中，已知 $\angle A : \angle B = 10 : 7$，$\angle B : \angle C = 7 : 3$，求这三个角，及三边的比率.
9. 在三角形 ABC 中，$AD \perp BC$，假定 $AB - AC = 3$，$BD - DC = 12$，且 $AD = 30$，求三边长.
10. 本题出自皮蒂斯楚斯《三角学之书》. 在多边形 $ABCDE$ 中，已知 $AB = 7$，$BC = 9$，$AC = 13$，$CD = 10$，$CE = 11$，$DE = 4$，$AE = 17$，且 $BF \perp AC$，$CG \perp AE$，$DH \perp CE$（图10.36），求多边形的面积.

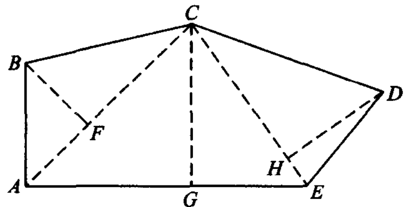


图10.36 皮蒂斯楚斯《三角学之书》的一个面积问题.

11. 本题出自哥白尼《天体运行论》. 已知等腰三角形的三边，求三个角. 作出三角形的外接圆，并以 A 点为中心，$AD = \frac{1}{2}AB$ 为半径作出另一圆（图10.37）. 由于两腰与底边之比等于半径 AD 与弦 DE 之比，所以三个角便可确定出来. 令 $AB = AC = 10$ 且 $BC = 6$ 来完成计算.
12. 证明 $\sin\alpha\sin\beta = \frac{1}{2}[\cos(\alpha - \beta) - \cos(\alpha + \beta)]$.
13. 用上题公式计算 4 378 218 乘以 $\sin 27°15'22''$. 并用计算器检验你的结果.

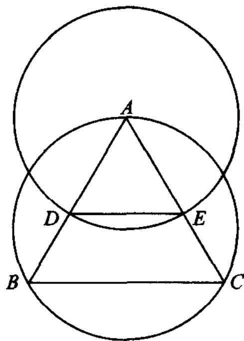


图10.37 求等腰三角形的角.

天文学问题

14. 已知地球的公转周期是1年，且火星的平均日距是地球的平均日距的1.524倍，应用开普勒第三定律确定火星的公转周期.
15. 按照开普勒第二定律，行星在其轨道上的哪一点公转最快？

对数问题

下面五道题通过稍微修改纳皮尔的定义，概括出一种方法来导出现代的自然对数.

16. 构想两条数字线，上面的线上标有一个算术序列 $\cdots, -4a, -3a, -2a, -a, 0, a, 2a, 3a, 4a, \cdots$，而下面的线上标有一个几何序列 $\cdots, 1/r^3, 1/r^2, 1/r, 1, r, r^2, r^3, \cdots$，其中 $r > 1$（图10.38）. 令 P 点在相同的时间内，越过各个相等的间隔 $[0, a]$，$[a, 2a]$，$[2a, 3a]$，$\cdots$，而 Q 点则在相同的时间内，越过各个递增的间隔 $[1, r]$，$[r, r^2]$，$[r^2, r^3]$，$\cdots$. 试说明定义中，Q 在任意所指出的点处的速度与该点到最左端的0点的距离成正比.

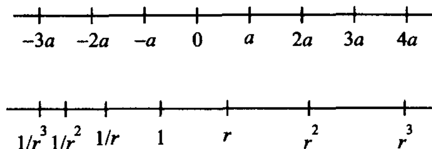


图10.38 现代对数概念的推导.

17. 试说明在下面的线上，只要 $\beta/\alpha = \delta/\gamma$，无论任意给定的区间 $[\alpha, \beta]$ 和 $[\delta, \gamma]$ 的长度如何，则 Q 点越过 $[\alpha, \beta]$ 和 $[\delta, \gamma]$ 所用的时间相等.
18. 在下面的线上，引入1与 r 之间的点 $s = \sqrt{r}$ 以及表示 s 的所有整数幂的点. 类似地，在上面的线上，引入点 $b = \frac{1}{2}a$ 以及表示 b 的整数倍的所有点. 试说明：如果仍然要求 Q 点在相同的时间内，越过各个新标出的间隔 $[1, s]$，$[s, s^2]$，$\cdots$，而 P 点以匀速连续地移动，那么习题16和17的结论仍成立. 这相当于在一定程度上缩小了下面线上

Q 点的速度的突变.

19. 为了完全排除速度的突变,假设 Q 点的速度是连续增加的,即下面线上任意点处的速度与该点到 0 点的距离成正比.再假设 P 点仍为匀速移动,速度为 v,并且 Q 点以速度 v 从 1 开始,而 P 点同时从 0 开始移动.令 Q 点位于数 x 处时,P 点同时位于数 y 处,则定义 y 是 x 的对数,记为 $\log x$.试说明 $\log 1 = 0$,并且当 $\beta/\alpha = \delta/\gamma$ 时,Q 点越过 $[\alpha,\beta]$ 和 $[\gamma,\delta]$ 所用时间相等.从而,有
$$\log\beta - \log\alpha = \log\delta - \log\gamma.$$
20. 从习题 19 的结论出发,试说明:
$$\log(\delta/\gamma) = \log\delta - \log\gamma,$$
$$\log(\beta\gamma) = \log\beta + \log\gamma,$$
$$\log(\beta^n) = n\log\beta, n \text{ 是整数},$$
$$\log(\sqrt[m]{\beta}) = \log\beta/m, m \text{ 是正整数},$$
$$\log(\beta^r) = r\log\beta, r \text{ 是有理数}.$$
21. 应用微积分说明,习题 19 定义的对数函数就是现代的自然对数.
22. 根据课本中纳皮尔对数函数 Nlog 的定义,用 Nlog x 和 Nlog y 表示 Nlog (xy) 和 Nlog (x/y).
23. 查索 17 世纪中,怎样通过计算尺的发明使得对数的计算和应用成为机械化.给出几种曾用过的计算尺例子.从何时起不再使用计算尺,为什么?

伽利略《两种新科学》中的问题

24. 证明:初速为零的物体沿斜面无摩擦下滑所用的时间与该物体从同样高度处自由降落的时间之比,等于斜面的长度与垂直高度之比.从而,沿相同高度的两个不同斜面下滑的时间之比,等于两斜面的长度之比.
25. 证明:初速为零的物体沿长度相同而斜度不同的两斜面下滑的时间之比,等于两斜面垂直高度之比的平方根的倒数.
26. 试说明以任意仰角 α 发射的炮弹,其轨迹是抛物线.
27. 以某一给定的初速度发射炮弹,当仰角 $\alpha = 45°$ 时,其水平跨度为 20 000.试说明当 $\alpha = 60°$ 或 $30°$ 时,以同样的初速度发射,它的水平跨度均为 17 318.
28. 以某一给定的初速度发射炮弹,当仰角 $\alpha = 45°$ 时,其最大高度为 5000.试说明若初速度不变,则当 $\alpha = 30°$ 和 $60°$ 时,该炮弹的最大高度分别为 2499 和 7502.
29. 已知初速为零的自由落体的位移与时间的平方成正比.证明在逐次相等的时间间隔内,位移为相邻的奇数 1,3,5,… 倍.

讨论题

30. 试比较伽利略和开普勒在新知识发展方面,对待实验(或观察)和理论之间的相互作用的态度.
31. 阅读本章参考书中所列出的 Arthur Koestler, *The Sleepwalkers*.评论该书作者的"梦游"假设在天文学新思想的发现方面有多大的可信性.
32. 查阅火星及其它行星轨道的偏心率.通过与地球偏心率的比较,想一想为什么开普勒要假设地球轨道的圆形性.分析这些偏心率,为什么开普勒详细研究的是火星而非水星?
33. 概括在习题 15 – 19 中的论述,对于在微积分之前引入自然对数,是否给出了一种合理的方法?并将这个方法与一般教材中的方法作以比较和评论.
34. 查阅现代绘画技巧中有关几何透视的论述,并与阿尔贝蒂的相关讨论作比较.
35. 查阅罗马天主教会有关伽利略一案的最新复审结论.教会是否修正了它认为伽利略不服从教会命令的观点?

文献和注解

下面列出几本详细讨论文艺复兴时期应用数学领域的书籍. Julian Lowell Coolidge, *The Mathematics of Great Amateurs*, Second Edition (Oxford: Clarendon Press, 1990) 含有三章内容论述文艺复兴时期艺术家们的数学研究,另有一章专论纳皮尔. Thomas S. Kuhn, *The Copernican Revolution*(Cambridge: Harvard University Press, 1957) 和 E.J. Dijksterhuis, *The Mechanization of the World Picture*(Princeton: Princeton University Press, 1986) 两本书都含有天文学发展的章节,后一书的数学过程更详细. Arthur Koestler, *The Sleepwalkers*(New York: Penguin, 1959) 一书对于从古希腊时期到伽利略时代的天文史给出了生动的描述,其中含有哥白尼,第谷,开普勒和伽利略的生平. Ernan McMullin, ed., *Galileo*, *Man of Science*(Princeton Junction: The Scholar's Bookshelf, 1988) 则是一本关于伽利略各方面科学成就的论文集.

1. John Dee, *Mathematical Preface*(New York: Science History Publications, 1975), Introduction. 这是一本影印本,无页码,附有 Allen G. Debus 关于迪伊生平及影响的论述. 关于迪伊的数学前言及其哲学的论述,亦见 F.A. Yates, *Theatre of the World*(Chicago: Chicago University Press, 1969).

2. Stillman Drake, *Galileo at Work*(Chicago: University of Chicago Press, 1978), pp. 347 – 348. 有大量作品论述伽利略及其与教会的冲突. 最有趣的卷本之一是 Pietro Redondi, *Galileo Heretic*, (Princeton: Princeton University Press, 1987).

3. John Dee, *Mathematical Preface*, p. 3.

4. 同上, p. 5.

5. 同上, p. 13.

6. 同上, p. 19.

7. 同上.

8. 同上, p. 38.

9. 同上.

10. 有关阿尔贝蒂的透视法成果,亦见 P. Green, "Alberti's Perspective: A Mathematical Comment," *Art Bulletin* 64(1987), 641-645.

11. 有关阿尔贝蒂和弗兰切斯卡的详细介绍,见 Julian Lowell Coolidge, *Mathematics of Great*, Chapter 3. 另外, J.V. Field, *The Invention of Infinity*: *Mathematics and Art in the Renaissance*(Oxford: Oxford University Press, 1997) 详细描述了文艺复兴时期艺术和数学的关系.

12. 该书的英译本,见 W. Strauss, trans., *The Painter's Manual*(New York: Abaris, 1977).

13. Erwin Panofsky. "Dürer as a Mathematician," in James R. Newman, ed., *The World of Mathematics*(New York: Simon and Schuster, 1956), vol. 1, 603 – 621, pp. 611 – 612. 这一章节引自作者本人论述迪勒生平及艺术成就的专著 Erwin Panofsky, *Albrecht Dürer*(Princeton: Princeton University Press, 1945).

14. 详见 Roger Herz-Fischler, "Dürer's Paradox or Why an Ellipse is not Egg-Shaped," *Mathematics Magazine* 63(1990), 75 – 85.

15. Dee, *Mathematical Preface*, p. 33.

16. 同上, p. 42.

17. 详见 Dava Sobel, *Longitude*: *The True Story of A Lone Genius Who Solved the Greatest Scientific Problem of His Time*(New York: Walker and Company, 1995).

18. Dee, *Mathematical Preface*, p. 15.

19. 详见 V. Frederick Rickey and Philip M. Tuchinsky, "An Application of Geography to Mathematics: History of the Integral of the Secant," *Mathematics Magazine* 53(1980), 162 – 166. 亦见 Florian Cajori, "On an Integration Ante-dating the Integral Calculus," *Bibliotheca Mathematica* (3)14(1914), 312 – 319.

20. Dee, *Mathematical Preface*, p. 20.

21. 同上,p.23.

22. Barnabas Hughes, *Regiomontanus on Triangles*(Madison: University of Wisconsin Press, 1967), p.27. 该书既有雷格蒙塔努斯《论各种三角形》的拉丁原文,又有英译,并附有一个介绍及大量注释.

23. 同上,p.101.

24. 同上,p.109.

25. 同上,p.119.

26. 有关贾比伊本艾夫拉赫与雷格蒙塔努斯著作之间的关系,详见 Richard Lorch in "Jābir ibn Aflah and the Establishment of Trigonometry in the West," in Richard Lorch, *Arabic Mathematical Sciences: Instruments, Texts, Transmission*(Aldershot, GB: Ashgate Publishing Limited, 1995).

27. 有关三角学的历史,详见 M.C.Zeller, "The Development of Trigonometry from Regiomontanus to Pitiscus"(Dissertation, University of Michigan, 1944, and Ann Arbor: Edwards Bros., 1946). 亦见 J.D.Bond, "The Development of Trigonometric Methods Down to the Close of the XVth Century," *Isis* 4(1921), 295 – 323.

28. A.M.Duncan, trans., *Copernicus: On the Revolutions of the Heavenly Spheres*(New York: Barnes and Noble, 1976), p.25. 这是哥白尼《天体运行论》的新英译本. 旧英译本是 Charles Glenn Wallis, appears in volume 16 of the *Great Books*(Chicago: Encyclopedia Britannica, 1952).

29. 同上,p.26.

30. 同上,p.27.

31. 同上,p.38.

32. 同上,p.60.

33. 同上,p.22.

34. A.M.Duncan, trans., *The Secret of the Universe*(New York: Abaris, 1981), p.67. 这是开普勒《宇宙的秘密》第二版的英译本,因此这不仅含有 1596 年的原文,而且含有开普勒在 1621 年所加的注释.

35. 同上,p.73.

36. Johannes Kepler, *Epitome of Copernican Astronomy*, Books IV and V, translated by Charles Glenn Wallis in the *Great Books*, vol.16, p.878.

37. Johannes Kepler, *The Harmonies of the World*, Book V, translated by Charles Glenn Wallis in the *Great Books*, vol.16, p.1048. 在钢琴上弹奏一下开普勒分配给各个行星的乐符,对于了解他所说的"天体的音乐"也许是一件有趣的事.

38. 关于开普勒对行星椭圆轨道的发现,详见 Curtis Wilson, "How Did Kepler Discover His First Two Laws," *Scientific American* 226(March 1972), 92 – 106, Eric Aiton, "How Kepler Discovered the elliptical orbit," *Mathematical Gazette* 59(1975), 250 – 260, and also Dijksterhuis, *Mechanization of the World*, pp.303 – 323.

39. Kepler, *Harmonies*, p.1020.

40. John Napier, *Constructio*, translated by William R.Mac-Donald (London: Dawsons, 1966), p.8.

41. Napier, *Constructio*, p.19.

42. John Napier, *Descriptio*, translated by Edward Wright(New York: Da Capo Press, 1969), p.6.

43. 有关纳皮尔对数表的建构方法,详见 C.H.Edwards, *The Historical Development of the Calculus*(New York: Springer-Verlag, 1979), Chapter 6.

44. Dee, *Mathematical Preface*, p.25.

45. 同上,p.34.

46. Galileo, *Two New Sciences*, translated by Stillman Drake(Madison: University of Wisconsin Press, 1974), p.154.

47. 详见 Stillman Drake, "Galileo's Discovery of the Law of Free Fall," *Scientific American* 228(May 1973), 84 – 92.

48. Galileo, *Two New Sciences*, p.160.

49. 同上,p.165.

50. Stillman Drake, ed., *Discoveries and Opinions of Galileo*(New York: Doubleday, 1957), p.183.

51. Galileo, *Two New Sciences*, p. 165.

52. 同上, p. 217.

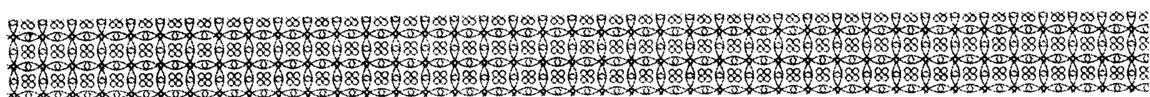
53. 同上, p. 196.

54. 详见 Stillman Drake, "Galileo's Discovery of the Parabolic Trajectory," *Scientific American* 232(March 1975), 102 – 110.

55. Galileo, *Two New Sciences*, p. 225.

文艺复兴时期应用数学概览

1377—1446	布努雷契(Filippo Brunelleschi)	透视学
1404—1472	阿尔贝蒂(Leon Battista Alberti)	透视学
1420—1492	弗兰切斯卡(Piero della Francesca)	透视学
1436—1476	雷格蒙塔努斯(Regiomontanus)	三角学
1471—1528	迪勒(Albrecht Dürer)	透视几何
1473—1543	哥白尼(Nicolaus Copernicus)	天文学
1502—1578	努涅斯(Pedro Nunes)	地图绘制
1512—1594	墨卡托(Gerard Mercator)	地图绘制
1514—1574	雷蒂库斯(George Rheticus)	三角学
1527—1608	迪伊(John Dee)	数学序言
1546—1601	第谷(Tycho Brahe)	天文学
1550—1617	纳皮尔(John Napier)	对数
1552—1632	比尔吉(Jobst Bürgi)	对数
1561—1613	皮蒂斯楚斯(Bartholomew Pitiscus)	三角学
1561—1615	赖特(Edward Wright)	地图绘制
1561—1631	布里格斯(Henry Briggs)	对数
1561—1656	芬克(Thomas Finck)	三角学
1564—1642	伽利略(Galileo Galilei)	运动学
1571—1630	开普勒(Johannes Kepler)	天文学



第11章 17世纪的几何、代数和概率

> 只要在最后的方程中出现两个未知量，我们就有一条轨迹，这两个量之一的末端描绘出一条直线或曲线.
>
> 《平面和立体轨迹引论》—— 皮埃尔·德·费马，1637[1].

1652年前后，安东尼·哥保德·德·默勒爵士(Antoine Gobaud)为了提高在赌博中获胜的机会，向布莱西·帕斯卡(Blaise Pascal)提出了两个关于赌博的问题.第一个是需要将两枚骰子掷多少次才能使得到两个6点的概率不小于50%；第二个问题是在赌博被打断时如何公平地分配赌注.概率论正是从帕斯卡对这两个问题的解答中诞生的.两人是通过罗内兹(Duke of Roannez)公爵认识的，在17世纪50年代早期，罗内兹公爵在巴黎有一个为数学家和另一些人提供碰头机会的沙龙.

17世纪早期，数学发展的步伐开始加快.当时，印刷工艺已经相当发达，通过书信和印刷品的通讯比从前大为快捷.一个数学家的想法更容易传达给其他数学家，供他们批评、评论并最终加以拓展.在这一章，我们考察一下几个刚发展起来的数学分支.韦达关于在分析中应用代数的想法在17世纪30年代在**解析几何**这一由代数和几何结合而来的学科中得到重新表述.解析几何发展中的两个中心人物是皮埃尔·德·费马(Pierre de Fermat)和勒内·笛卡儿(René Descartes)，而解析几何的这一发展在随后的微积分的发明中是至关紧要的.这两个人在数学的另一些领域也扮演了主要角色.笛卡儿和托马斯·哈略特(Thomas Harriot)和阿尔伯特·吉拉德(Albert Girard)一起将韦达的一些代数思想发展成方程理论.费马在同布莱西·帕斯卡的通信中参与了概率论的早期发展，这一理论的第一本教科书是在1656年由克里斯蒂·惠更斯(Christian Huygens)撰写的.费马还是自斐波那契(Leonardo of Pisa)以来第一个撰写数论著作的人，而帕斯卡则同德扎格(Girard Desargues)一起对射影几何作出了最初的贡献.

11.1 解析几何

诞生于1637年的解析几何有两个创始人，勒内·笛卡儿(1596—1650)和皮埃尔·德·费马(1601—1665).当然此前存在一个酝酿时期，但在1637年初，费马将一份题为《平面和立体轨迹引

论》的手稿寄给了他在巴黎的一个朋友. 大约在同一时候,笛卡儿的《更好地指导推理和寻求科学真理的方法论》已经完成了校对,正准备付梓,该书有三个附录,其中之一就是《几何学》. 不论是费马的《引论》还是笛卡儿的《几何学》都表现出将代数学和几何学联系起来的同一些基本技巧,这些技巧后来在近代的解析几何学科中发展到登峰造极的地步. 费马和笛卡儿都是在试图重新发现古希腊人"失传"的分析技巧的努力中找到这些技巧的. 两个人都非常熟悉古希腊经典文献,尤其是帕普斯的《分析学科》,并且两个人还都是用阿波罗尼乌斯的四线轨迹问题及其推广作为他们新思想的试金石的. 但是费马和笛卡儿对同一个学科作了迥然不同的处理,他们的区别植根于他们对数学本身的不同见解上.

11.1.1 费马和《平面和立体轨迹引论》

费马的数学学习开始于在图卢兹听到的师范大学课程,这些课程很可能比欧几里得几何的入门课程多不了多少内容. 但他在取得学士学位后开始法律学习前的时间里,在波尔多跟随韦达的几个学生学习了好几年数学,韦达的这几个学生在 17 世纪 20 年代后期正致力于编辑出版他们老师的著作. 费马因此熟悉了韦达关于代数符号化的新思想以及他发现和阐明希腊数学家神秘的分析学的计划. 在波尔多,费马开始了借助帕普斯在《分析荟萃》中的注释和引理来重构阿波罗尼乌斯的《平面轨迹》的计划. 费马试图复原阿波罗尼乌斯的原著和他在发现各个定理时所作的解释. 费马对韦达的研究很自然地使他试图用一个代数版本取代阿波罗尼乌斯的几何分析. 正是阿波罗尼乌斯轨迹定理的代数版本构成了费马解析几何的开端.

例如,费马考虑了以下结果:如果从任意多个给定点向一点引直线,以得到的线段形成的正方形面积之和等于给定的面积,则该点位于一个确定的圆周上. 这一定理涉及到不定数目个点,但费马对只有两个点的最简单情形的处理却包含了解析几何两个主要思想的精粹,几何轨迹和包含两个或多个未知量的不定方程的对应以及建立这一对应的几何框架 —— 一个包括轴和长度的系统 —— 都是明白无误的.

人物小传

皮埃尔 · 德 · 费马(1601—1665)(Pierre De Fermat)

费马生于法国南部博芒特 – 德 – 洛马格内的一个殷实的家庭里,他的父亲是一个皮革商人,还是一名当地的低级官员. 他在图卢兹大学接受了本科教育并在 1631 年在奥尔良取得民法学士学位. 然后他回到图卢兹并终生在这附近从事法律工作. 他是图卢兹包括议会内阁在内的许多官方团体的成员,议会内阁兼有行政和司法职能. 虽然费马当了许多年法律工作者,显然他从来都不是一名出色的律师,也许是因为他花费了大量时间在他的最爱 —— 数学 —— 上. 因为他的健康状况和法律工作的压力,他从未游历过远离家乡的地方. 因此,他所有的数学工作都是通过他广泛的通信传达给别人的.

费马一直把数学当成是一种嗜好,是使他从作为法律工作者不得不处理的持续的纷争中解脱出来的避难所. 所以他拒绝发表他的任何发现,因为这样做会迫使他完成每一细节并可能把他卷入另一领域的纷争中. 在许多情况下,即使存在证明,人们也不知道费马提出的是什么证明,人们还不知道他的某部分工作是否存在系统的叙述. 费马经常通过暗示他解决某些问题的新方法来挑逗他是通信伙伴. 他有时会勾画出这些方法的轮廓,但他"如果有空"就填补空白的承诺却经常得不到兑现. 无论如何,对他死后 14 年由他儿子出版的手稿和许多书信的研究使得学者们如今对费马的方法有了比较完善的印象.[2]

费马选取了两个给定点 A 和 B，E 是 AB 的中点.他以 IE 为直径（I 仍有待确定），E 为圆心画圆（图 11.1）[3].接着他说明了只要选择 I 使得 $2(AE^2+IE^2)=M$，圆上的任意一点 P 均满足定理的条件，即 AP^2+BP^2 等于给定的面积 M.这个证明的重要思想作为轨迹的圆由两个变量 AP 和 BP 的平方和决定，而点 I 由从"原点"E 测量的"坐标" 确定.

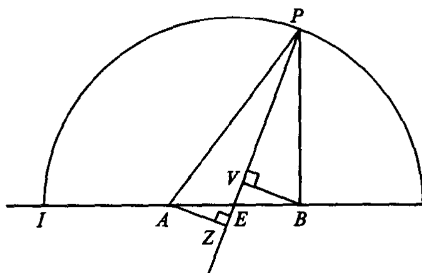


图 11.1 费马对阿波罗尼乌斯定理的一个特殊情形的分析.

$$
\begin{aligned}
AP^2+BP^2 &= PZ^2+AZ^2+PV^2+BV^2\\
&=(PE+EZ)^2+AE^2-EZ^2\\
&\quad+(PE-EV)^2+BE^2-EV^2\\
&=PE^2+2PE\cdot EZ+EZ^2+AE^2\\
&\quad-EZ^2+PE^2-2PE\cdot EV+EV^2\\
&\quad+BE^2-EV^2\\
&=2PE^2+2AE^2\\
&=2(AE^2+IE^2).
\end{aligned}
$$

利用一个原点来决定一个点的（水平）坐标的思想在费马对涉及几个不共线的点的阿波罗尼乌斯定理的处理中也是显而易见的.对于那种情况，费马选择一条基准线，使得所有给定的点都位于该基准线的一侧，然后从每个点向基准线作垂线.在他对问题的分析中，他不但用到了水平坐标 GH、GL、GK，他还用到了沿基准线垂线度量的纵坐标 GA、HB、LD 和 KC（图 11.2）.实际上，他证明了所求圆的中心 O 的水平坐标 GM 由 $GM=0.25(GH+GL+GK)$ 决定，纵坐标 $MO=0.25(GA+HB+LD+KC)$.半径 OP 由方程 $M=AO^2+BO^2+CO^2+DO^2+4OP^2$ 决定，M 是给定的面积.

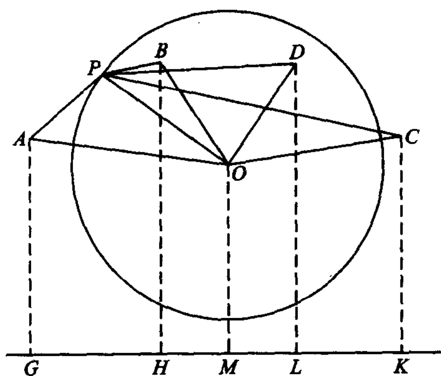


图 11.2 阿波罗尼乌斯定理的第二种特殊情形.

费马在对阿波罗尼乌斯定理的一般情形进行处理的时候，却没有用方程来描述所求的圆，可能他是力求用阿波罗尼乌斯的风格来写作.但他在完成重构工作的两年之后，他在他的《平面和立体轨迹引论》中记下了他关于解析几何的新思想，本章开头引用的那句话更表达了它的中心思想.费马断定，如果在用代数方法求解几何问题时，最终得到的是一个有两个未知量的方程，得到的解就是一个轨迹，或者是直线或者是曲线，轨迹上的点由一个变动的线段端点的运动决定，另一个端点沿一条固定的直线运动.

费马在这一简要的介绍中要表达的主要观点是，如果一条运动的直线同一固定直线成一固定夹角，并且两个未知量的最高次不超过 2，则得到的轨迹将是一直线、圆或者另外的圆锥曲线.他进而依照出现的不同情形逐一证明了他的结果.我们首先对直线的情形作一说明："设 NZM 是一位置

固定的直线，N 是固定点. 令 NZ 等于未知量 A，形成角 NZI 的线段 ZI 代表另一未知量 E. 如果 D 倍的 A 等于 B 倍的 E，点 I 将表示位置固定的一条直线."[4] 费马从一个单轴 NZM 和一个线性方程开始(图 11.3)(费马沿用了韦达用元音字母表示未知量，辅音表示已知量的约定). 他希望证明我们现在写作 $dx = by$ 的方程表示一条直线. 因为 $D \cdot A = B \cdot E$，也就是 $B : D = A : E$，因为 $B : D$ 是一个已知的比例，所以比值 $A : E$ 也就确定了，三角形 NZI 也因此确定. 因此直线 NI 的位置就是确定的. 费马省略掉了完成论证所必需的一个"容易"的步骤，说明 NI 上任意一点 T 依比例 $NW : TW = B : D$.

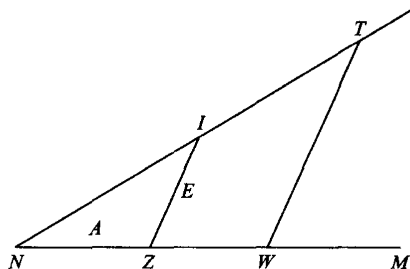


图 11.3　费马对 $D \cdot A = B \cdot E$ 的分析.

虽然近代解析几何的基本概念在费马的描述中是清楚的，费马的思想同我们现在的思想有所不同. 第一，费马只使用一个轴. 曲线不是被理解成由相对于两条轴画出的点组成的，而是由变动的线段 ZI 的端点 I 当另一端点 Z 沿给定的轴运动时产生的. 费马经常取 ZI 和 ZN 的夹角为直角，虽然这样做没有什么特别的必要. 第二个区别是对费马、韦达或者当时的大多数其他人而言，代数方程唯一适合的解为正解. 因此费马的"坐标" ZN 和 ZI，作为方程 $D \cdot A = B \cdot E$ 的解，代表正数. 所以，费马只画出了从原点到第一象限的一条射线.

费马第一象限的局限在他对抛物线的处理中也清楚可见："如果 Aq 等于 D 倍的 E，则点 I 在一抛物线上."[5] 费马希望证明方程 $x^2 = dy$(用现代的符号表示)确定一个抛物线. 他从两条基本的线段 NZ 和 ZI 开始，在本例中它们成直角. 画 NP 平行于 ZI，他接着断定以 N 为顶点、NP 为轴、D 为正焦弦的抛物线由给定的方程确定(图 11.4). 费马当然假定他的读者对阿波罗尼乌斯的《圆锥曲线》非常熟悉. 对抛物线的情形，阿波罗尼乌斯的画图说明了由 D 和 NP 构成的矩形和 PI(或 NZ)为边的正方形面积相等，该命题翻译成代数语言就是 $dy = x^2$. 虽然费马知道抛物线是什么样的形状，他的图形只包括了抛物线一半的一部分. 他没有涉及沿轴线的负的长度.

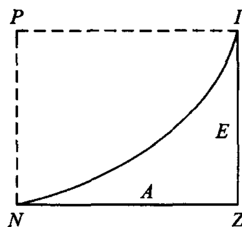


图 11.4　费马对方程 $Aq = D \cdot E$ 的分析.

费马进而去确定由其它 5 种包含两个变量的二次方程所代表的曲线. 用现代的符号表示，$xy = b$ 和 $b^2 + x^2 = ay^2$ 表示双曲线，$b^2 - x^2 = y^2$ 表示圆，$b^2 - x^2 = ay^2$ 表示椭圆，而 $x^2 \pm xy = ay^2$ 表示直线. 在每种情况下，他的论证采用按阿波罗尼乌斯的方法构造圆锥曲线，进而说明该圆锥曲线具有给定的方程. 最后，费马勾画了通过变量变换将任意二次方程化为他的七种标准形式之一的方法. 例如，他断定，如果轴与形成曲线的夹角为直角，任何包含 ax^2 和 ay^2 以及 bx 和(或) cy 的方程都可以化为圆的标准形式. 所以，方程 $p^2 - 2hx - x^2 = y^2 + 2ky$ 可以通过在方程的两边加上 k^2 使得右边变成完全平方来变换. 令 $r^2 = h^2 + k^2 + p^2$ 或者 $r^2 - h^2 = k^2 + p^2$，费马可以把原方程重写为 $r^2 - (x + h)^2 = (y + k)^2$，如果把 $x + h$ 替换为 x'，$y + k$ 替换为 y'，则成为圆的标准方程. 费马还通过变量变换的方法处理了包含 xy 项的方程.

费马能够确定对应于任意包含两个变量的二次方程的轨迹，并说明该轨迹一定是直线、圆或者圆锥曲线. 在《引论》的结尾，费马注意到人们可以将他的方法应用于四线轨迹问题的如下推广："如果从同一个点向任意多个给定直线引成已知角度的直线，如果所得的线段的平方和等于给定的面积，则该点位于一位置确定的圆锥曲线上."[6] 但是费马将实际的解答留给了读者.

11.1.2 笛卡儿和《几何学》

费马简短的论文在它到达巴黎时引起了震动.以玛林·梅森(Marin Mersenne,1588—1648)为中心的一伙数学家们经常在一起聚会讨论数学和物理中的新思想.梅森的作用像是这个团体中负责记录和联络的秘书,他以此角色从各个来源接收资料,复制这些资料然后广为传播.因此梅森就像是法国的"活的科学杂志".费马从1636年开始和梅森建立起了经常的通信联系,但是因为费马的许多手稿都很简洁并且缺少细节,梅森经常要求费马扩展他的论著.无论如何,对《引论》的反响是积极的,这为费马奠定了第一流数学家的声誉.费马的手稿传到巴黎并被笛卡儿看到的时间刚好是在笛卡儿出版他自己的解析几何著作之前.人们可以想象得出,当笛卡儿在他的作品发表之前就看到了类似的东西时会有多么懊恼.

人物小传

勒内·笛卡儿(1596—1650)(René Descartes)

笛卡儿(图11.5)生于图尔附近的拉-哈耶(如今的拉-哈耶-笛卡儿)的一个旧的法国贵族家庭.因为他在整个青年时代体弱多病,他在学生时代被允许晚起.他因此养成了在早晨沉思的习惯.他的思考使他得出结论,他在学校所学的东西几乎都是不确定的.事实上,他变得如此多疑以至于他决定弃学.如同他在《方法论》中所记述的,"我利用我青年时代剩下的时间去旅游,去看宫殿和军队,去频繁造访有着不同性情和生活条件的人们,去积累各种经验,去在命中注定的际遇中考验自己,并随地思索发生的事情,这样我从这些事情中就能有所收益."[7]例如,他参加了三十年战争中的几个战役,之后在1628年他定居荷兰,在那里开始用毕生的时间去创造一种适用于发现关于世界的真理的哲学.他决定只接受那些如此清晰明了以至不会引起任何怀疑的思想作为真理,然后模仿数学的推理模式通过简单合乎逻辑的步骤去发现新的真理.他很快撰写了一篇物理学的重要论文,但在最后关头,他听说了教会对伽里略的宣判,因此他决定不去发表这篇论文以防一处小的教义上的错误导致他的整个哲学遭到禁止.但他很快相信他应该与世人分享他的新思想.在1637年,他发表了《方法论》,该书附带有说明他"方法"效力的关于光学、气象学和几何学的三篇论文.

随着另外几部哲学著作的发表,笛卡儿的国际声誉得以提高,1649年他受到瑞典女王克里斯蒂娜的邀请去斯德哥尔摩作她的家庭教师.他不情愿地接受了.不幸的是,他的健康抵抗不了北方气候的严酷,特别是因为克里斯蒂娜要求他早起,这同他长年养成的习惯相反.不久笛卡儿就患了肺病,这导致了1650年他的死亡.

然而,笛卡儿的解析几何同费马的还是有所不同的.要理解这点,人们必须首先认识到,《几何学》的写作目的是为了展示笛卡儿在《方法论》中讨论的建立在自明原理基础上的正确推理的方法在几何学中的应用.同费马一样,笛卡儿研究过韦达的著作并从中找到了理解希腊人分析方法的钥匙.但笛卡儿更关心的是通过用几何方法构造代数方程的解来说明代数和几何的关系,而不是通过研究轨迹来说明这种关系的.那么在某种意义上,他不过是追随了一个古老的传统,在他之前诸如海亚姆和图西这样的伊斯兰数学家曾经继续过这个传统.但笛卡儿的确像费马一样迈出了关键的一步,这是他的伊斯兰先辈们从没有迈出过的一步,即应用坐标来研究几何和代数的关系.



图11.5 法国邮票上的笛卡儿和他的《方法论》.

《几何学》是这样开始的:"任何一个几何问题都很容易化归为用一些术语来表示,使得只要知

道直线段的长度的有关知识，就足以完成它的作图.”[8] 在这一包含三卷的著作的第一卷中，笛卡儿解决了仅需要直线和圆这些标准欧几里得曲线的几何问题. 但因为毫不隐晦地使用了代数技巧，笛卡儿使得这些欧几里得技巧看起来非常现代. 比如说，为了作出二次方程 $z^2 = az + b^2$ 的解，他作直角三角形 NLM，使得 $LM = b$，$LN = 0.5a$（图 11.6）. 延长斜边到 O，使得 $NO = NL$，以 N 为圆心，NO 为半径作圆，他的结论是 OM 就是所要求的 z，因为 z 的值由以下标准公式给出：

$$z = \frac{1}{2}a + \sqrt{\frac{1}{4}a^2 + b^2}.$$

在同样的条件下，MP 是方程 $z^2 = -az + b^2$ 的解，并且，如果作 LN 的平行线 MQR，则 MQ 以及 MR 是方程 $z^2 = az - b^2$ 的两个解.

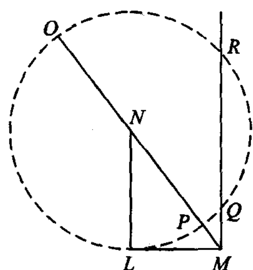


图 11.6　笛卡儿的解二次方程的作图.

不过，笛卡儿注意到：“通常，我们并不需要在纸上画出这些线，而只要用单个字母来标记每一条线段就够了.”[9] 只要我们知道了哪些运算在几何上是可能的，我们就可以只进行代数演算并把结果表述为公式. 在这些代数运算中，笛卡儿迈出了关键的另外一步. 笛卡儿用 a^2 和 a^3 这样的项表示线段而不是几何学上的正方形和立方体 —— 因此笛卡儿是第一个一贯地采用了现代的幂符号的人. 所以，笛卡儿还可以考虑更高次的幂而无须担心它们没有几何意义. 笛卡儿曾提到任何表达式都可以被当作包含了单位的任意次方幂，这就是他对韦达严格遵守的齐性要求仅有的一点考虑，但实际上笛卡儿自由地对代数表达式求和而根本不管它们的次数. 并且，笛卡儿用字母表末尾的字母表示未知数而用字母表开头的字母表示已知数，笛卡儿用这一目前仍通行的作法取代了韦达的作法，韦达采用元音和辅音字母表示未知数和已知数.

在第一卷的结尾，笛卡儿详细地讨论了阿波罗尼乌斯的四线问题. 正是在这里笛卡儿引进了一个坐标轴作为所有直线和所求轨迹的参照系. 问题要求找到所有的那些点，当从这些点向四直线以一定角度引直线时所得到的两条线段的积与另两条线段的积之比为定值. 笛卡儿注意到：“因为总有无穷多个不同的点满足这些要求，所以需要发现和描绘出含有所有这些点的曲线.”[10]

利用图 11.7，笛卡儿发现如果所有直线都以两条主直线为参照，问题便可以得到简化. 因此，他设 x 为线段 AB 沿已知直线 EG 的长度，y 为线段 BC 沿所求线段 BC 的长度，C 是满足问题要求的一点. 符合题意的线段 CB、CH、CF 和 CD（由点 C 分别引向给定的直线 EG、TH、FS 和 DR）的长度都可以表示为 x 和 y 的线性函数. 例如，因为三角形 ARB 的所有角都为已知，比值 $BR : AB = b$ 也就已知，由此可得 $BR = bx$，$CR = y + bx$. 因为三角形 DRC 的三个角也为已知，所以比值 $CD : CR =$

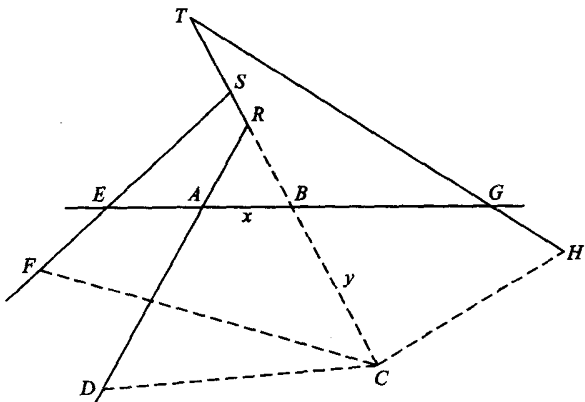


图 11.7　四线轨迹问题.

c 也已知，故 $CD = cy + bcx$. 同理，令固定的距离 $AE = k, AG = l$，已知比值 $BS : BE = d, CF : CS = e, BT : BG = f, CH : TC = g$，人们依次可以得到 $BE = k + x, BS = dk + dx, CS = y + dk + dx, CF = ey + dek + dex, BG = l - x, BT = fl - fk, CT = y + fl - fx$ 以及 $CH = gy + fgl - fgx$. 因为问题涉及到某些对长度乘积的比较，所以表示所求轨迹的方程是 x 和 y 的二次方程，并且可以画出任意多所求轨迹上的点. 因为任给一个 y 值，x 的值就被一个确定的二次方程表示出来，该方程的解已经给出. 所求的曲线可以由此绘出. 在《几何学》的第二卷，笛卡儿重新回到这一问题并且证明了两个变量的二次方程表示的曲线或者是圆或者是一种圆锥曲线，这取决于方程中各个常系数的取值.

笛卡儿在《几何学》中主要关心的是作为几何问题解的那些点的实际作图. 对这一工作来说，确定这种作图中哪些类型的曲线是可以接受的就很必要. 自然，他用来作出四直线问题轨迹上的点的圆和直线属于合法曲线并且是最简单的合法曲线. 圆和直线也是欧几里得用过的曲线. 但其他的希腊著作家却毫不迟疑地使用了圆锥曲线和某些另外的曲线.[11]

笛卡儿决定把他的几何学中可以接受的曲线的定义建立在欧几里得画直线和圆的公理 1、公理 3 和以下一个新公理的基础上，"两条或者两条以上的直线可以一条在另一条之上移动，并由它们的交点确定出其它曲线".[12] 因此他只接受由某些机械描绘出的曲线. 究竟哪些曲线符合笛卡儿的定义在今天并不是十分清楚，但笛卡儿曾经给出过若干专门被设计出来绘制曲线的仪器的例子. 例如，在图 11.8 中，GL 是以 G 为枢轴的直尺. 它在点 L 同装置 $CNKL$ 连接在一起，装置允许 L 沿 AB 移动并始终保持直线 KN 同其自身平行. 两条移动的直线 GK 和 KN 的交点 C 确定一条直线. 通过简单的几何学考虑，笛卡儿找到了这一曲线的方程. 令 $CB = y, BA = x$，并且设常量 $GA = a, KL = b, NL = c$，笛卡儿计算出 $BK = (b/c)y$, $BL = (b/c)y - b$ 以及 $AL = x + (b/c)y - b$. 因为 $CB : BL = GA : AL$，笛卡儿得到以下方程

$$\frac{ab}{c}y - ab = xy + \frac{b}{c}y^2 - by,$$

或者，最终地，

$$y^2 = cy - \frac{c}{b}xy + ay - ac.$$

笛卡儿不加证明地断定，该曲线是双曲线.

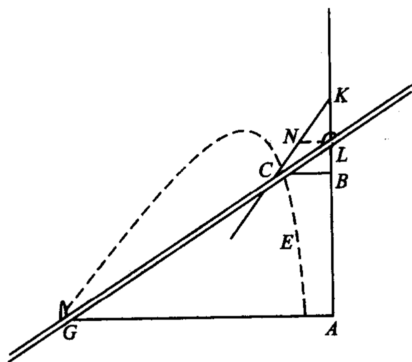


图 11.8 笛卡儿的绘制曲线的仪器.

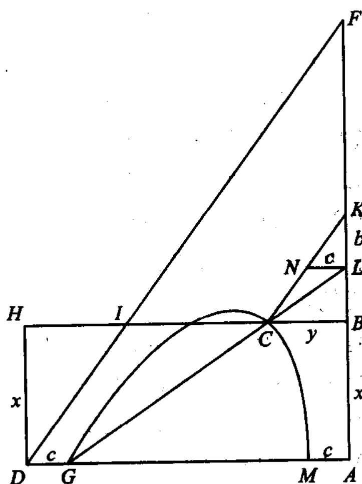


图 11.9 范·舒滕对笛卡儿双曲线作图的证明.

据推测,笛卡儿指望他的读者懂得足够多的阿波罗尼乌斯的著作以至于可以理解为什么由该装置绘出的曲线是双曲线.然而,弗兰斯·范·舒滕(Frans van Schooten,1615—1660)在他 1649 年的评注中感觉有必要给出一个证明.延长 AG 到 D,使得 $DG = MA$(图 11.9).因为点 M 是当 GL 同 GA 重合时得到的曲线上的一点,因此 $MA = NL$.现在作 DF 平行于 NK,交 AK 于 F(我们也可以把 DF 看作是当 GL 平行于 KN 时 KN 的延长).现在我们可以通过点 M 作一个以 DF 和 AF 为渐近线的双曲线,我们将证明该双曲线同已经作出的双曲线是同一个.延长 BC 交 DF 于 I,作 DH 平行于 AF 交 BC 于 H.有三角形 KLN 相似于三角形 DIH,所以 $KL : NL = DH : HI$.但是 $DH = AB = x$,故 $HI = cx/B$,$IB = AD - HI = AG + DG - HI = a + c - cx/b$,所以 $IC = IB - BC = a + c - cx/b - y$.根据阿波罗尼乌斯的《圆锥曲线》的 Ⅱ - 10(见第 3 章),$IC \cdot BC = DM \cdot MA$.所以,

$$\left(a + c - \frac{cx}{b} - y\right)y = ac \quad 或者 \quad y^2 = cy - \frac{c}{b}xy + ay - ac,$$

所作双曲线同原来的双曲线有相同的方程.

笛卡儿定义他所说的“几何的”曲线的基本理由似乎是“这些曲线上的所有的点,必定跟直线上的所有的点具有一种确定的关系,而且这种关系必须用单个的方程来表示”.[13] 换句话说,任何一个这类曲线必须能表示为一个代数方程.显然,笛卡儿也相信这个命题的反命题,任何两个变量的代数方程是一个曲线,该曲线的作图能用适当的机械来实现.他无法证明这一命题,但他在《几何学》第三卷中投入了大量的篇幅来说明如何作二次以上代数曲线上的点.笛卡儿相信那些人们能作出任意点的曲线都可以用他的某一个机器的连续运动来描绘.

笛卡儿为什么用连续运动定义“几何的”曲线而不是直接定义成有一个代数方程的曲线可能有几种原因.第一,笛卡儿感兴趣的是几何学研究的改革.用纯粹的代数标准来定义可以接受的曲线会把他的工作归结为代数学.第二,因为他希望能作出作为几何问题的解的点,他需要能确定代数曲线的交点.用连续运动定义曲线会直接确定出交点对笛卡儿来说是显而易见的.用代数方程定义的曲线是否有交点则不是很清楚.依照他的基本哲学,笛卡儿不会把一个代数定义接受为公理.最后,笛卡儿显然不相信代数方程是定义一个曲线的最好方式.《几何学》中没有一处是从方程开始的.与费马不同,笛卡儿总是先用几何描述一个曲线,然后在适当的时候才推导出它的方程.方程对笛卡儿来说仅仅是一种研究曲线的工具而不是定义标准.

另一方面,为什么笛卡儿排斥那些无法几何地加以定义的曲线的确是个问题.他当然知道没有代数方程的曲线.一个古老的例子是割圆曲线,它由一个转动和直线运动的组合加以定义(见图 3.6).这一曲线之所以能像困扰古人那样困扰笛卡儿是因为两种运动不存在精确的可以度量的关系,原因是人们无法精确地确定圆周长和半径的比率.正如笛卡儿所说:“由于我们并不知道直线和曲线之间的比,而且我相信这种比是人的智力所无法发现的.因此,不可能基于这种比而得出任何严格和精确的结论.”[14] 对笛卡儿来说不幸的是,17 世纪 50 年代对各种曲线精确长度的首度确定以及对非几何曲线(或者说超越曲线)下面积的研究很快就瓦解了笛卡儿对几何上能接受和不能接受曲线的根本区分.

很清楚,费马和笛卡儿两人都明白几何曲线和两个变量的代数方程的基本联系.两人都没有使用今天所用的两个坐标轴而是把单个坐标轴作为他们的基本工具,沿着这个坐标轴对两个未知量中的一个进行度量,两个人也都没有坚持度量第二个未知量的直线要同单个坐标轴垂直.虽然两人都能够作出二次以上的方程的曲线,但他们都使用熟悉的圆锥曲线作为主要范例.最后,虽然两人处理的都是曲线而不是函数的解析几何,但每个人都以其自己的方式理解了函数的思想,即一个变量的变化决定了另一个的变化.

但是两人是从不同的角度来研究解析几何的课题的.费马明确提出过一个两个变量的方程确定一条曲线.他总是从方程开始,然后才绘出曲线.与此不同,笛卡儿更关心几何.对他来说,曲线是第一位的.给定曲线的几何描述,他能够列出方程.因此笛卡儿不得不处理比费马处理过的方程远为复杂的代数方程.正是笛卡儿的方程的复杂性使他发现了处理高次多项式方程的方法,这一方法将在11.2.2节中得到讨论.

笛卡儿和费马强调了方程和曲线关系的两个不同方面.遗憾的是,费马从未出版过他的著作.虽然费马的工作表述得非常清楚并且以手稿的形式在欧洲传播,它却达不到出版著作的影响.与此相反,笛卡儿的著作非常难读.它用法语而不是通常的拉丁语出版,在推理上跳跃很多,包含大量的复杂方程,这使得数学家们很少能完全理解它.笛卡儿实际上对书中的跳跃很是自得.他在著作的结尾写道:"我希望后世会给予我仁厚的评判,不单是因为我对许多事情作出的解释,而且也因为我有意省略了的内容——那是留给他人享受发明之愉悦的".[15]但在《几何学》出版若干年后,笛卡儿的态度有所改变.他鼓励其他的数学家将他的著作译成希腊文并发表解释他的思路的评论.只是在莱顿大学的工程学教授范·舒滕的拉丁文版本出版后,笛卡儿的著作才获得了他所渴望的承认.这一最早出版于1649年的拉丁文版本带有范·舒滕本人和弗劳里蒙德·德波那(1601—1652)的注释,随后在1659—1661年增加了更多的注释和扩充.

11.1.3 让·德·威特的工作

对范·舒滕1659—1661版的笛卡儿《几何学》作出的补充之一是让·德·威特(1623—1672)的一篇关于圆锥曲线的专著.在学生时代,德·威特曾师从范·舒滕,而范·舒滕则认识笛卡儿并且在巴黎逗留期间曾经研究过费马的著作.通过范·舒滕,德·威特熟悉了解析几何全部两个发明者的著作.在1646年德·威特23岁的时候,他编著了《曲线基础》一书,书中他从综合和分析两个角度处理了圆锥曲线的题目.两卷本的《基础》的第一卷致力于应用综合几何的传统方法推导各种圆锥曲线的性质.第二卷是使用新方法论述圆锥曲线的第一本系统专著,这里德·威特从两个变量的方程出发将费马的思想扩展成为对圆锥曲线的完全的代数处理.虽然德·威特的方法论同费马的相似,但他采用了笛卡儿现代的符号.



图11.10 荷兰邮票上的德·威特.

人物小传	让·德·威特(1623—1672)(Jan De Witt)
	德·威特是一名有才华的数学家,但他在数学上几乎不能投入什么时间.他生于一个政治上活跃的荷兰家庭并成为了他家乡多特镇的首长.在奥兰治家族的威廉二世国王死后,他在1653年当选为荷兰的大州长,因此实际上成为总理.他带领荷兰度过了以后19年的艰难岁月,成功地平衡了英国和法国相矛盾的要求.但当法国在1672年入侵荷兰时,人民要求威廉三世重掌政权,随后发生了反对德·威特的暴力游行,他被一伙狂怒的暴徒所杀害.

例如,定理Ⅰ说方程 $y = bx/a$ 的轨迹是一条直线.德·威特的证明同费马的相似,但他利用相似性明确说明他所作直线上的任意一点的坐标 x, y 都满足关系 $a : b = x : y$.像费马一样,德·威特在常量和坐标上都仅仅考虑了正值,因此所求直线看起来仅是一条从原点出发的射线.但他进一步说明另一些方程也确定直线:

$$y = \frac{bx}{a} + c,\ y = \frac{bx}{a} - c,\ y = c - \frac{bx}{a}, y = c \text{ 以及 } x = c.$$

但是,在所有的情况下,都只有落入第一象限的那部分直线被画出.

同样,德·威特也像费马一样地去证明 $y^2 = ax$ 代表一条抛物线.他也画出了由方程 $y^2 = ax + b^2, y^2 = ax - b^2, y^2 = b^2 - ax$ 以及通过将其中 x, y 互换形成的方程决定的抛物线的图像.像从前一样,只有当 x, y 均为正值时的那部分图像被画出.但德·威特还仔细考虑了更复杂的方程:

$$y^2 + \frac{2bxy}{a} + 2cy = bx - \frac{b^2x^2}{a^2} - c^2.$$

令 $z = y + bx/a + c$ 可将方程化为 $z^2 = \frac{2bc}{a}x + bx$,或者再设 $d = \frac{2bc}{a} + b$ 方程化为 $z^2 = dx$,德·威特知道这是一个抛物线的方程.他然后说明如何利用这个变换绘出轨迹.如果点 D 的坐标为 (x, y),以 AE 为 x 轴,AF 为 y 轴,他令 $BE = AG = c$ 并延长 DB 到 C 使得 $GB : BC = a : b$ 或者 $BC = bx/a$ (图 11.11).可以推出 $DC = y + c + bx/a = z$.同样,令 $GB : GC = a : e$ 得到 $GC = ex/a$.用现代的术语,德·威特应用变换 $x = (a/e)x'; y = z - (b/e)x' - c$ 完成了从斜轴 AE 和 AF 到垂直轴 GC 和 GF 的转换;因而具有坐标 (x, y) 的点 D 通过变换获得了同原来的坐标相关的新坐标 (x', z).在新的坐标下,曲线的方程现在成了 $z^2 = (da/e)x'$,这是一个顶点为 G,轴为 GC,正焦弦长为 da/e 的抛物线.德·威特从而画出了这个抛物线,或者更具体地说,是原来坐标轴 AE 以上的部分 ID,因为只有这部分可以作为所求的轨迹.注意到给定轨迹上的任意一点 D,抛物线的基本性质意味着 DC 上正方形的面积等于以 $GC (= ex/a)$ 和正焦弦 da/e 为边的矩形的面积,德·威特完成了证明.因此 $z^2 = dx$,且由代换 $z = y + c + bx/a$ 可重新得到原方程.

图 11.11　德·威特对 $y^2 + \frac{2bxy}{a} + 2cy = bx - \frac{b^2x^2}{a^2} - c^2$ 的构图.

德·威特以类似的方法详细地处理了椭圆和双曲线,他首先给出了诸如

$$\frac{ey^2}{g} = f^2 - x^2 \quad (\text{椭圆}),$$

以及

$$\frac{ey^2}{g} = x^2 - f^2 \text{ 和 } xy = f^2 \quad (\text{双曲线})$$

的标准形式,然后说明如何通过适当的代换把其它的方程化成标准形式之一.虽然德·威特没有表述出决定轨迹是抛物线、椭圆还是双曲线的原始方程的条件,通过分析他的例子发现这些是很容易的.在德·威特著作的最后,他总结出任何一个两个变量的二次方程都可以化为标准形式中的一个,因此代表了直线、圆或者圆锥曲线.虽然费马和笛卡儿也大体得到了同样的结果,但给出了解决二次方程轨迹问题所有细节的人仍然是德·威特.

11.2　方程理论

解三次和四次方程的代数方法发现于 16 世纪的意大利,并在接近 16、17 世纪之交时由韦达作出了某些改进.但是卡尔达诺受到了缺乏方便符号的阻碍而韦达则始终局限于正数解.因此尽管前

者用各种例子给出了单个三次方程根之间的关系以及相关联的方程的根之间的关系，而后者能够在所有值为正的条件下给出直到5次的方程的系数和解之间关系的代数表达，方程的一般理论仍然是不完整的.

11.2.1 托马斯·哈略特和他的数学手稿

托马斯·哈略特(1560—1621)仔细研究过韦达的工作，并且因为他理解处理方程负根甚至虚根的必要性，他在通往一般理论的方向上获得了进展.哈略特在牛津完成了他的本科学习，随后参加了瓦尔特·拉雷爵士的军队并且作为制图专家在1585年参加了到弗吉尼亚的探险.在那里他除了学会了吸烟，这个习惯最终使他死于癌症，他还编写了一份关于殖民地以及当地居民的简要报告.

哈略特的大部分数学工作至今都还是手稿，虽然许多手稿似乎曾在他生前和死后在英国流传.但是因为他从不花时间把他的工作写成适合发表的形式，他的许多思想都以未完成的形式保留在笔记里，其《分析艺术的实践》在他死后发表于1631年，它仅仅考虑了方程的正根.

哈略特继承了韦达用元音字母表示未知量而用辅音字母表示已知量的思想，不过他用小写字母代替了韦达的大写字母，但他用重复的单个字母改进了韦达幂的符号.比如他把 a^4 写成 $aaaa$，而不是用"平方－平方"的一个简写.他也认识到方程可以将表达式 $b-a, c-a, d-a, \cdots$ 相乘从它们的根 $b, c, d, \cdots$ 产生.由此他得到了甚至包括负根和虚根情形的方程根和系数的基本关系，不过他似乎从未将这个结果明确表述成定理.

例如，哈略特注意到如果将 $b+a, c-a$ 以及 $df-aa$ 相乘，就可以得到他记为

$$bcdf - bdf\,a - df\,aa + b\,aaa + cdf\,a - bc\,aa - c\,aaa + aaaa = 0000$$

的方程，他还注意到方程的根为 $a=c, a=-b$ 以及 $a=\sqrt{df}$.在这个例子中，他似乎没有意识到 df 应该有两个平方根，但在下面的例子中，他甚至可以得出方程的两个复根.在方程 $12 = 8a - 13aa + 8aaa - aaaa$ 中，他首先看到2和6是两个根.注意到这些实根之和已经等于 a^3 的系数8，他断定不可能再有其它的实根，因为它们将使得和大于8.(观察他的前一个例子，我们会奇怪他是如何得到这一结论的.)但他进而用代换的方法求解方程.令 $a=2-e$，他得到新的方程 $-20e+11ee-eeee=0$，方程的实根是0和－4.另两个根的和一定是4，而它们的积则为5.哈略特接着写出了另两个根是 $e=2+\sqrt{-1}$ 以及 $e=2-\sqrt{-1}$，因而原方程的复根为 $a=2-(2+\sqrt{-1})=-\sqrt{-1}$ 和 $a=2-(2-\sqrt{-1})=+\sqrt{-1}$.[16]

11.2.2 阿尔伯特·吉拉德和代数基本定理

阿尔伯特·吉拉德(1595—1632)在他1629年的著作《代数学的新发现》中对多项式根和系数的关系比哈略特更为清楚，并且他还给出了代数基本定理的第一个明确表述.吉拉德可能出生在法国罗瑞纳省的圣米希尔，但他一生的大部分时间在荷兰度过，他曾在莱顿读大学并在纳塞公国的弗里德里克·亨利的部队里当过军事工程师.虽然他写过三角学的书并编辑过斯蒂文的著作，他最重要的贡献是在代数学方面.在《代数学的新发现》中，吉拉德明确引进了分数指数，"分子是幂而分母是根"，[17] 他还引进了现在表示高次根的记号，例如表示3次方根的 $\sqrt[3]{\ }$ 就是指数1/3的另一种表达.并且，他还是最早注意到方程负数解几何意义的人之一："负数解在几何上解释为倒退；负数向后退而正数向前进."[18] 他甚至还给出了一个几何问题的例子，它的代数翻译有两个正解和两个负解，他在相关的图上注意到负数解可以解释成在正数解相反的方向上放置.

吉拉德不仅理解方程负数解的意义，他还把韦达和哈略特的工作系统化并且明确考虑了今天

称为 n 个变量的基本对称函数的**集团**(faction):"给出若干个数,其总和可以被称作第一集团;全部两两相乘的积之和被称为第二集团;全部三三相乘的积之和被称为第三集团,这样一直进行到最后,但所有数的积称为最后集团.现在,集团的数量和数的总数一样多."[19] 他指出,对2,4,5来说,第一集团是11,为它们的和;第二集团是38,两两相乘的积的和.第三集团是40,全部三个数的积.他还注意到他称为"抽取三角"的二项式系数的帕斯卡三角显示每个集团包含多少项.在4个数的情况,第一集团包含4项,第二集团6项,第三集团4项,第四也是最后集团1项.

吉拉德在代数理论中的基本结果是以下定理,他没有提供定理的证明:

定理 每一个代数方程容许有同方程的次数同样多的解.而且解的第一集团等于次高项的系数,解的第二集团等于第三高次项的系数,第三集团等于第四高次项的系数,以此类推,所以最后集团等于常数项——所有这些在符号上可以用交替的顺序来记录.[20]

吉拉德关于符号的最后一句话的意思,是人们首先需要将方程加以组织使得次数交替出现在方程的两边.例如 $x^4 = 4x^3 + 7x^2 - 34x - 24$ 应该被写成 $x^4 - 7x^2 - 24 = 4x^3 - 34x$.这个方程的根是1,2,-3和4,第一集团等于4,x^3 的系数;第二集团等于 -7,x^2 的系数;第三集团等于 -34,x 的系数;第四集团等于 -24,常数项.同样方程 $x^3 = 167x - 26$ 可以改写为 $x^3 - 167x = 0x^2 - 26$.因为 -13 是一个解,他的结果意味着其余两个根的积是2,而它们的和为13.要找到它们只需要求解一个二次方程.答案为 $6.25 + \sqrt{40.25}$ 以及 $6.25 - \sqrt{40.25}$.

在定理的第一部分,吉拉德断言代数基本定理的真实性,每个多项式方程解的个数与它的次数相等.如同他的例子所显示的,他承认一个给定的解可能出现一次以上.他还充分认识到在统计解的个数时他必须把虚根包括在内(他称虚根为不可能的).所以在他的例子 $x^4 + 3 = 4x$ 中,他注意到四个集团为0,0,4,3.因为1是一个二重根,剩余两个根具有积为3和为 -2 的性质.可以得出它们是 $-1 \pm \sqrt{-2}$.在回答这些不可能解的价值这个意料中的问题时,吉拉德说:"它们有三方面的好处:在一般法则的确定性方面,在确定没有其它的解方面以及在它的效用方面."[21]

吉拉德没有解释不可能解的"效用"是什么.他也没有说明他是怎么推导出这一定理的.即使他考虑过多重解,但他似乎像哈略特那样把 n 次方程理解成通过 n 个表达式 $x - r_i$ 相乘得到的,而某一些 r_i 也许是相等的.不过使这一过程精确化的却是笛卡儿.

11.2.3 笛卡儿和解方程

一旦知道了一个解 α,吉拉德在某些场合利用他的集团来降低方程的次数.给原多项式除以 $x - \alpha$ 的标准方法最早是由笛卡儿在他的《几何学》第三卷中加以阐明的.笛卡儿通过引用——几乎是——吉拉德的结果开始了他自己对方程的研究,"每一个方程都可能有跟方程中未知量的次数一样多的不同的根."[22] 笛卡儿使用"可能有"代替了吉拉德的"容许有",因为他只考虑不同的根而且至少在最初他不愿意考虑虚根.但在后来他确实注意到有时候根为虚值以及"我们总可以想象,每一个方程都具有我已指定的那么多根,就是说同次数相等,但并不总是存在确定的量同如此想象得到的每个根相对应."[23]

笛卡儿明确说明了由方程的解构造方程的方法.例如如果 $x = 2$ 或者 $x - 2 = 0$ 并且同时 $x = 3$ 或者 $x - 3 = 0$,笛卡儿注意到两个方程的积是 $x^2 - 5x + 6 = 0$,这是一个根为2和3的二次方程.同样,如果这后一个方程同 $x - 4 = 0$ 相乘,就得到一个以2,3,4为根的3次方程,$x^3 - 9x^2 + 26x - 24 = 0$.进而同具有"假"根5的方程 $x + 5 = 0$ 相乘得到一个具有4个根的4次方程,3个"真"根和1个"假"根.笛卡儿作结论说:"显然,由上述讨论可知,具有若干个根的方程的各项之和,即多项式

本身,总能被这样的二项式除尽,它由未知量减去真根之一的值或加上假根之一的值组成.据此,方程的次数可以被降低.反过来,若方程各项的和不能被由未知量加或减某个别的量组成的二项式除尽,则这后一个量不是方程的根."[24] 这是现代因式定理的最早表述.同他平时的风格一致,笛卡儿没有给出完整的证明.他仅仅说这个结果是"显然的".

同样,笛卡儿还不加证明地表述了今天人们称为**笛卡儿符号法则**的结果:"一个方程的真根数目跟它所含符号的变化、即从 + 到 - 或从 - 到 + 的多寡一致;而其假根的数目,跟连续找到两个 + 号或两个 - 号的次数一样."[25] 例如,方程 $x^4-4x^3-19x^2+106x-120=0$ 有 3 次符号变化和 1 对连续的 - 号,所有它可能有的正根最多为 3 个,负根最多为 1 个.实际上,方程的根为 2,3,4 以及 -5.

但笛卡儿更关心的是方程解的作图,所以在接近第三卷的末尾时他明确示范了一些高次方程的作图方法.特别是对于3次和4次方程,他使用抛物线同圆相交的方法,抛物线和圆都符合他可作曲线的标准.笛卡儿的方法同海亚姆的方法相似,但与他的伊斯兰的前辈不同,笛卡儿意识到某一些交点表示方程的负(假)根以及"若圆跟抛物线既不相交也不相切,这表明方程既无真根也无假根,此时所有的根都是虚的."[26]

笛卡儿进而说明了如何利用圆和用他的某种机械画出的曲线相交的方法求解 4 次以上的方程.虽然他只简要地勾画了他的方法并在若干例子中有所运用,但笛卡儿相信"对于复杂程度越来越高的问题,我们只要遵循同样的、具有普遍性的方法,就能完成其作图;因为对数学级数来说,一旦得到了最前面的两项或三项,其余的便很容易得到."[27] 直到17世纪末,许多数学家努力把笛卡儿的方法加以推广,用几何方法去求解各种各样的方程.但是几何方法对于透彻地理解方程解的性质是不够的.人们发现即使对求解笛卡儿展示过他的作图技巧的那些几何问题而言,代数方法以及微积分的新思想也更加适合.

11.3 初等概率论

人们通常认为现代的概率理论开始于帕斯卡和费马在 1654 年的通信,这些通信在某种程度上是为了回应德·默勒给帕斯卡提出的赌博问题,这点在本章开篇时有过提及.因为赌博是最古老的休闲活动之一,所以很可能人们从远古时代就考虑过概率的基本概念,至少是在经验的基础上,特别是人们至少拥有在某一赌局中如何计算特定事件发生的可能性的模糊的观念.骰子在若干古老的文明中已经被发现.虽然人们并非总能确定这些骰子的用途,但它们非常可能被用于预测未来以及用于赌博.遗憾的是,关于如何进行各式各样的赌博以及是否存在关于可能性的任何计算,在这些文明中没有任何文字资料保存下来.

在可以上溯到纪元之初的犹太文献中,关于这些计算人们知道稍微多一些的情况,虽然这些是有关各种犹太法律的应用而不是关于赌博的.塔木德经是一部记录着拉比们在犹太法律解释方面所作的讨论的著作,它包含着在确定由已知概率事件复合而成的事件的概率时加法定律和乘法定律的应用.这些概率当时被用来评判各种决定.虽然有证据表明人们能直觉地理解一些概率定律并把它们应用在决策理论上,没有迹象说明存在过详尽系统的工作.[28]

11.3.1 概率论的最初开端

中世纪晚期欧洲出现了同骰子游戏有关的一些基本的概率论思想.例如,由若干计算两枚或三枚骰子可以掷出的不同方式数的记录,两枚骰子时为 21 种,3 枚时是 56 种.假定我们只计算可能出现的不同点的组合而不考虑这些点出现的顺序,这些数字是正确的.例如,在两枚骰子的情形,只有

一种方式掷出 2 点，一种方式掷出 3 点，两种方式掷出 4 点(2,2 和 3,1)，两种方式掷出 5 点(1,4 和 2,3)，诸如此类．用现代的说法，这些方式不是“等可能的”，因而不能用作计算胜算的依据．但是计算骰子可能掷出的方式最可能出自更早时期的骰子在占卜中的应用，在占卜中骰子掷出的实际结果决定未来而并不牵扯到什么胜负比．现存最早的关于 3 枚骰子可以掷出的 56 种方式不是等可能的记述出现在一首作者不详的拉丁诗歌《维图拉》(De Vetula) 中，这首诗大约完成于 1200—1400 年间：“如果 3 枚骰子点数一样，对每个点数就只有一种方式；如果有 2 枚骰子点数一样而另一枚不一样，则有 3 种方式；如果 3 枚都不一样就有 6 种方式．”[29] 根据提到过的规则对具体情况的分析表明 3 枚骰子可以掷出的总的方式数为 56 种(图 11.12)．

Tabula II.

Omninò Similes.					
666	555	444	333	222	111
Duo Similes et tertius diſsimilis.					
665	664	663	662	661	
556	554	553	552	551	
446	445	443	442	441	
336	335	334	332	331	
226	225	224	223	221	
116	115	114	113	112	
Omninò Diſsimiles continui.					
654	543	432	321		
Discontinui.					
642	531	641	631		
Duo Continui et tertius discontinuus.					
653	652	651	621	521	421
542	541	643	431	632	532

19

Quinquaginta modis & ſex diverſificantur
In punctaturis, punctaturæque ducentis
Atque bis octo cadendi ſchematibus, quibus inter
Compoſitos numeros, quibus eſt luſoribus uſus,
Diviſis, prout inter eos ſunt diſtribuendi,
Plenè cognosces, quantæ virtutis eorum
Quilibet esſe poteſt, ſeu quantæ debilitatis:
Quod ſubſcripta poteſt tibi declarare figura.

Tabula III.

Qvot Punctaturas, et qvot Cadentias habeat qvilibet numerorũ compoſitorum.

3	18	Punctatura	1	Cadentia	1
4	17	Punctatura	1	Cadentiæ	3
5	16	Punctaturæ	2	Cadentiæ	6
6	15	Punctaturæ	3	Cadentiæ	10
7	14	Punctaturæ	4	Cadentiæ	15
8	13	Punctaturæ	5	Cadentiæ	21
9	12	Punctaturæ	6	Cadentiæ	25
10	11	Punctaturæ	6	Cadentiæ	27

图 11.12　《维图拉》中的一页，它表示了三个骰子可能掷出的所有 56 种方式(来源：Houghton 图书馆，哈佛大学)．

到 16 世纪，等可能事件的思想开始被理解，因而进行实际的概率计算成为可能．进行这种计算的系统尝试最早出现在卡尔达诺在 1526 年撰写的《论机会游戏》中，尽管该书在他生前并未发表．卡尔达诺在准确地计算出两枚或者三枚骰子能够掷出的方式数之外，还显示出对概率基本概念的理解．例如，卡尔达诺首先统计出掷两枚骰子出现一个 1 点的不同方式共 11 种，出现 2 点的另有 9 种，出现 3 点的此外还有 7 种，然后计算出对于要求掷出 1,2 或 3 点的问题，有 27 种成功的投掷和 9 种不成功的投掷，因此胜负比是 3 : 1．由此推出公平的赌注应该是赌出现 1,2 或者 3 的人出 3 枚硬币，赌不出现的出 1 枚硬币．因为在 4 次投掷中他们可能会取得平手．

卡尔达诺还意识到了对独立事件的概率的乘法法则，但他在他的著作中记录了他最初对究竟用哪些量相乘存在的困惑．例如，他计算出投掷 3 枚骰子至少出现一个 1 点的机会是 216 次中的 91 次，所以赌反方的胜负比是 125 : 91．为了确定在连续两次投掷中不出现 1 个 1 点的胜负比，他将原先的胜负比平方，计算的结果是 15625 : 8281，或者说接近 2 : 1．但经过一番考虑，他发觉这种推理必

错无疑,因为如果一个给定事件的机会是50%(胜负比是1:1),这种推理将意味着这一事件连续出现两次或三次的机会仍然是50%.他说,这是“最荒谬不过的,因为如果一个使用两枚骰子的赌徒掷出偶数和奇数的机会相等,这并不说明他在连续三次投掷中都掷出偶数的机会也是50%.”[30] 卡尔达诺随后改正了错误.在对一些简单情形作了仔细计算之后,他意识到相乘的应该是概率而不是胜负比.例如,一个赌局的胜负比为3:1,或者说成功的概率是3/4,卡尔达诺证明了在连续的两局中,接连成功的机会为9次,不能接连成功的机会是7次.因此,两次成功的概率是9/16,而胜出的机会为9:7.他进而作出推广,发现将总结果数为f,成功数为s的试验重复n次,正确的胜出机会是s^n比$f^n - s^n$.

卡尔达诺还讨论过德·默勒曾经向帕斯卡提出的问题,即确定掷一对骰子允许掷多少次才能使出现两个6点的机会达到50%,显然该问题已经流行多年了.卡尔达诺的观点是因为每36次中有1次机会掷出两个6点,所以平均起来每掷36次这样的结果会出现一次.因此,在半数这么多次投掷即18次投掷中,出现两个6点的机会会达到50%.在处理一枚骰子的情况时他作了同样的推理,2点在3次投掷中出现的机会是50%.卡尔达诺的推理意味着在6次投掷中出现一次2点是一定的,在36次投掷两枚骰子时,出现两个6点是一定的,但他没有意识到他的错误.

德·默勒向帕斯卡提出的关于赌注分配的问题此前也在意大利被考虑过,尤其是在卢卡·帕乔里的《摘要》中.帕乔里版的问题是有两个人在进行一场公平的赌博,赌局将在一个人赢过6轮后结束.赌博实际在一个人赢5轮另一个人赢3轮时中断.帕乔里对赌注分配问题的解答是赌注应该按5:3的比例分配.塔尔塔利亚在他完成于大约60年后的《论数字与度量》(Generale Trattato)一书中注意到这个答案肯定是错误的,因为帕乔里的推理意味着如果游戏停止时一个参加者赢了一局而另外一个人赢了零局,第一个人将拿走全部赌注,这显然是不公平的结果.塔尔塔利亚争辩说因为两个得分相差两局,第一个人需要赢的次数仅是第二人的1/3,第一个人应该拿走第二个人赌注份额中的1/3,所以总的赌注应该按照2:1的比例分配.塔尔塔利亚显然对他的答案也没有十分的把握,因为他在结论中说:“这样一个问题的解决是法律上的而非数学上的,所以无论怎样分配都有理由上诉.”[31]

11.3.2 布莱西·帕斯卡,概率和帕斯卡三角

卡尔达诺和塔尔塔利亚关于概率的思想没有被他们同时代的数学家所接受,而是被遗忘.直到1660年前后的10年,概率才进入了欧洲思想,事情是在两种意义上发生的,首先作为理解在偶然性的过程中稳定频率的方式,其次还作为确定信任的合理程度的方法.布莱西·帕斯卡(1623—1662)的工作例示了所有这两种意义.在他对德·默勒分配问题的数学解答中,帕斯卡在同机会博弈打交道,但在他为信仰上帝所作的决策论式的辩护中,却没有任何关于机会的概念.

帕斯卡在1654年给费马的几封信中描述了他对分配问题的解法,若干年后他在《论算术三角形》的末尾又作了更为详细的描述.他从适用于分配的两条基本原理开始.第一,如果一个给定的参与人的处境是不论他赢或者输,某一数额都归他所有,则即便赌博中断时他也应该得到这一数额.第二,如果两人的处境是,若一人赢,则某一数额归他;如果此人输,则该数额归对方,并且假定赢的机会均等,则他们在无法进行赌博时应该均分这一数额.

帕斯卡继而注意到决定赌注分割的是剩下的总局数以及依照规则每个参与人为获得全部赌注需要赢的局数.因此,如果他们在两胜制的赌博中战成1:0,或者在3胜制的赌博中战成2:1,或者在11胜制的赌博中战成10:9,在中断时赌注分配的结果应该是一样的.在任何一个场合,第一个参

与人都需要再赢一场，而第二人则需要两场.

> **人物小传**
>
> **布莱西 · 帕斯卡(1623—1662)(Blaise Pascal)**
>
> 帕斯卡(图 11.13)生于法国的克勒蒙特 – 费兰，他很早就显示了数学方面的早熟. 他的父亲埃提纳在他年轻时就把他介绍到梅森周围的圈子里. 因此，年轻的帕斯卡很快就熟悉了包括费马工作在内的法国数学的主要进展. 他在 20 岁以前就开始了自己的数学和科学研究. 他的成就包括有一种计算机器的发明以及对大气压下流体行为的研究. 但在 1654 年以后，他在科学上的兴趣被在宗教事务上不断增长的兴趣所压倒. 他的健康状况一直不佳，39 岁时死于一场暴病.

作为帕斯卡原理的一个示例，假定角逐的总赌注是$80. 首先，假定每个人都需要再赢一场时比赛中断，简单地将$80 一分为二. 第二，假定第一人需要赢一场而第二人需要赢两场. 如果第一人赢得了下一场比赛，他将赢得$80. 如果他输了，则两人都只需要赢一场，所以根据第一种情况，第一人将赢得$40. 如果他们此时停止比赛，第一人将有权得到他无论输赢都会得到的$40 加上剩余$40 的一半，即$60，亦即两次可能赢取数量的平均数. 同样，假如第一人需要赢一场而第二人需要赢 3 场，下一局有两种可能性. 如果第一人获胜，他赢得$80，而如果他输掉，则处境同上一例相同，当时他有权得到$60. 由此推出如果下一场没有进行，第一人将得到$60 加上剩余$20 的一半，即$70，亦即两次可能赢得数量的平均数.



图 11.13　法国邮票上的帕斯卡.

结果表明分配问题的通解需要帕斯卡三角的某些性质. 在考察帕斯卡的解答之前，我们必须先看看他对他称之为算术三角形的对象的构造和运用，该数字三角形在世界的各个地区已经被使用了 500 年以上. 帕斯卡的《论算术三角形》—— 该书还以对数学归纳法的明确表述而闻名 —— 从他对三角形的构造开始，三角形的左上角从 1 开始，再使用每个数等于它上面的数和左边的数之和的

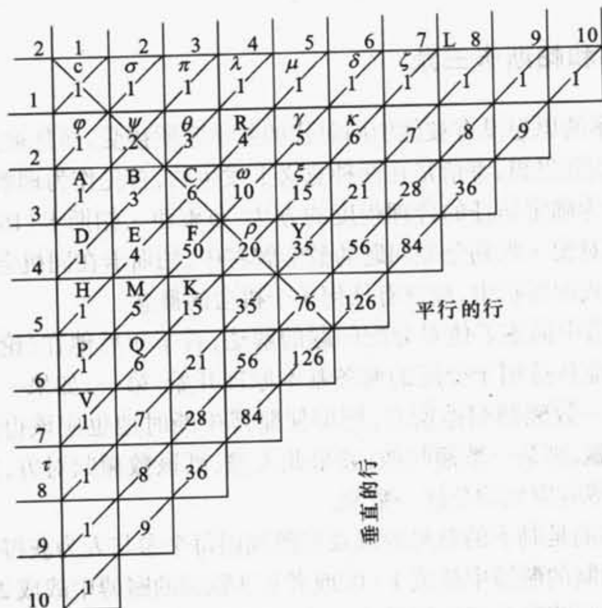


图 11.14　算术三角形的帕斯卡形式.

法则来构造(图11.14).但是,在讨论帕斯卡的结果时,使用现代的表格和现代的符号来识别三角形的各个项将会更为清晰.标准的二项式符号$\binom{n}{k}$将被用来命名第n行的第k项(开始的行和列编号都为0).基本的构造法则这时成为

$$\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1}.$$

行 \ 列	0	1	2	3	4	5	6	7	8
0	1								
1	1	1							
2	1	2	1						
3	1	3	3	1					
4	1	4	6	4	1				
5	1	5	10	10	5	1			
6	1	6	15	20	15	6	1		
7	1	7	21	35	35	21	7	1	
8	1	8	28	56	70	56	28	8	1

帕斯卡从考虑各个项如何同其它项的和相联系开始了他的研究.他的证明通常采用"可以推广的例子"的方法,因为像他的前辈一样,他也没有将通项符号化的好的方式.例如,帕斯卡的"第三个结果"(关于三角形的定义)断定每一项都等于它前面一列截至到它前面一行的所有元素之和

$$\binom{n}{k}=\sum_{j=k-1}^{n-1}\binom{j}{k-1}.$$

在这一场合,帕斯卡选取特殊项$\binom{4}{2}$为例,根据构造方法,该项等于$\binom{3}{1}+\binom{3}{2}$.因为

$$\binom{3}{2}=\binom{2}{1}+\binom{2}{2}\text{以及}\binom{2}{2}=\binom{1}{1},$$

结果成立.

帕斯卡对第8个结果的证明使用了数学归纳法,该结果说,第n行元素的和等于2^n,从k到$k+1$归纳的一步在第7个结果中得以实现:任何一行元素的和是前一行元素和的一倍.这一命题的证明还是采用了可以推广的例子的方法.他取了一个特殊行,第三行,注意到第一项和最后一项等于第二行的第一和最后一项,而第三行的其它各项等于第二行两项之和.因此,第三行项的和就包含了第二行中的每个元素两次.通过简单地指出第0行仅包含一个元素1,它的和等于2^0,而每个后继行都是前一行的2倍,帕斯卡完成了第8个结果的证明.

奇怪的是,仅仅是在第12个结果的证明中,帕斯卡才明确表述了数学归纳法的原理,还不是以最一般的形式而仅仅是在要证明的特定结果的背景下表述的.该结果为

$$\binom{n}{k}:\binom{n}{k+1}=k+1:n-k.$$

帕斯卡写到:"虽然这一命题有无穷多种情况,我将通过引进两条引理来简要地证明它,"两条引理即是归纳论证的两个基本部分."第一个,是自明的,这个比例在第一行成立,因为$\binom{1}{0}:\binom{1}{1}=1:1$

是十分明显的.第二个是如果这个比例对任意一行成立,则它在下一行也成立.由此它显然也一定在所有行成立.因为根据第一条引理,它在第二行成立;所以根据第二条引理它也在第三行成立,所以第四行也成立,以至无穷."[32] 虽然这对于手头的这个特殊例子来说是归纳法原理的清楚陈述,对应用它证明一般结果的理由的陈述也很清楚,但帕斯卡仍然没有一般地证明第二条引理而只是证明了引理在第三行成立意味着在第四行也成立.例如,为了证明$\binom{4}{1}:\binom{4}{2}=2:3$,他首先发现$\binom{3}{0}:\binom{3}{1}=1:3$,因此

$$\binom{4}{1}:\binom{3}{1}=\left(\binom{3}{1}+\binom{3}{0}\right):\binom{3}{1}=4:3.$$

然后,因为$\binom{3}{1}:\binom{3}{2}=2:2$,可以推出

$$\binom{4}{2}:\binom{3}{1}=\left(\binom{3}{2}+\binom{3}{1}\right):\binom{3}{1}=4:2.$$

用第一个比例式除以第二个比例式就可得要求的结果.帕斯卡意识到这一证明不是普遍的,因为他补充作了以下说明:"对于所有其它行证明是一样的,因为它只需要找出上一行的比例式以及每一项等于它上面一项和左边一项的和,这是普遍成立的."[33] 无论如何,这第 12 项结果使得帕斯卡通过合比简单地证明出:

$$\binom{n}{k}:\binom{n}{0}=(n-k+1)(n-k+2)\cdots n:k(k-1)\cdots 1.$$

或者因为$\binom{n}{0}=1$,也就是:

$$\binom{n}{k}=\frac{n(n-1)\cdots(n-k+1)}{k!}.$$

在完成了对算术三角形基本性质的叙述后,帕斯卡说明了如何在若干领域将它付诸应用.他应用归纳法推理证明了$\binom{n}{k}$等于在n个元素的集合中取出k个的组合数.他还证明了三角形的行元素是二项式系数,就是说,n行的数字是$(a+1)^n$的展开式的a的各次幂的系数.但帕斯卡相信,三角形的一个更为重要的应用是在赌注分配问题方面.他如下解决了这一问题

定理　假定第一个参与人缺r场最后获胜,第二个参与人缺s场,r和s不小于 1.如果整场角逐就此停止,赌注应当如此分配使得第一个参与人得到全部赌注的比例为$\sum_{k=0}^{s-1}\binom{n}{k}:2^n$,这里$n=r+s-1$(剩余局数的最大值).

定理断定第一个参与人获胜的概率是二项展开式$(1+1)^n$的前s项的和同总和2^n之比.我们可以把展开式的第一项理解为第一个参与人赢得n点的机会数,第二项是赢得$n-1$点的机会数,以此类推,第s项给出了赢得$n-(s-1)=r$点的机会数.因为我们也可以假定实际上刚好需要再进行n场比赛,这些系数给出了所有第一个参与人获胜的方式.

帕斯卡从$n=1$或者说$r=s=1$的情形开始用归纳法证明了定理,在这种情况赌注应该被均分.定理断言赌注的分配应该使第一个参与人得到比例为$\binom{1}{0}:2$,或者说 1/2,所以结果对$n=1$成

立.下一步是假定结果在最多剩余局数为 m 时成立,证明在最多剩余局数为 $m+1$ 的情况下成立,这时第一个参与人缺 r 局而第二个参与人缺 s 局.同以往一样,帕斯卡对这一归纳步骤的证明使用了可推广的例子,取 $m=3$.但我们将应用现代符号给出完全的证明.考虑如果参与人需要再玩一场时的两种可能性.如果第一个参与人获胜,他将缺 $r-1$ 场而第二个参与人仍然缺 s 场.因为 $r-1+s-1=m$,归纳假设表明第一个参与人应该得到的赌注的比例是 $\sum_{k=0}^{s-1}\binom{m}{k}:2^m$.另一方面,如果第一个参与人输掉了下一场比赛,归纳假设表明他应该得到的赌注的比例是 $\sum_{k=0}^{s-2}\binom{m}{k}:2^m$.因此,根据帕斯卡的基本原理,当下一场比赛没有进行时,第一个参与人应该得到那两个值的平均值,即得到赌注的比例是:

$$\sum_{k=0}^{s-1}\binom{m}{k}+\sum_{k=0}^{s-2}\binom{m}{k}:2\cdot 2^m.$$

二项式系数的和可以重写为

$$\binom{m}{0}+\sum_{k=1}^{s-1}\binom{m}{k}+\sum_{k=1}^{s-1}\binom{m}{k-1}.$$

根据算术三角形的构造法则,并因为 $\binom{m}{0}=\binom{m+1}{0}$,这一和式最终等于 $\sum_{k=0}^{s-1}\binom{m+1}{k}$.因为 $2\cdot 2^m=2^{m+1}$, 对 $n=m+1$ 的情形,第一个参与人之所得正如定理所断言,证明得以完成.

帕斯卡因此彻底回答了德·默勒的分配问题.在他与费马的通信中,两人讨论了在参与人多于两个时的这一问题,两人取得一致的解答.帕斯卡还简要地提及了另一个问题,确定投掷两枚骰子出现两个6点的机会达到50%需要的次数.他注意到在投掷一枚骰子的类似问题中,四次投掷出现一个6点的胜负比是671比625,但没有说明他计算出这一结果的方法.德·默勒显然相信,因为投掷1枚骰子时(此时有6种可能的结果),4次投掷足以保证至少对等的胜负比,所以不论投掷多少枚骰子,4:6的比例应该保持不变.因为投掷两枚骰子有36种可能的结果,他认为正确的值应当是24.他也许因为这一数值在经验上并不正确才向帕斯卡提出了这一问题.帕斯卡发现在24次投掷中胜负比将小于50%,但他在信件中或者其它地方并没有对这一论断背后的理论详加说明.

帕斯卡支持对上帝的信仰的决策论式的论证展示了概率推理的第二个方面,达到"合理"决策的方法.按照帕斯卡的观点,上帝或者存在或者不存在.对两个命题哪个为真一个人只能进行"押宝",这里押的宝就是他的行动.换句话说,一个人或者在行动上毫不在意上帝或者在行动上遵循(基督教的)上帝的概念.一个人应当如何行动?如果上帝不存在,怎么行动就没多大差别.但如果上帝存在,押上帝不存在将受到诅咒而押上帝存在将得到拯救.因为后者的结局比前者要好无穷倍,这一决策问题的结论就很清楚,即使一个人相信上帝存在的可能性不大:"理性"的人还是要象假定上帝存在那样行动.

11.3.3 克里斯蒂·惠更斯和最早的概率论课本



图11.15 荷兰邮票上的惠更斯.

根据帕斯卡的前提,他支持对上帝信仰的论证当然是成立的.(一个人是否接受他的前提是另一回事.)实际上,他的以某种形式计算一个特定行动的"价值"的概念成为了概率论方面第一篇系统论文的基础,该论文是范·舒滕的一个学生克里斯蒂·惠更斯(1629—1695)(图11.15)在1656年写成的.惠更斯在1655年一次对巴黎的访问过程中对概率的问题产生了兴趣并写下了关于这一课题的

一本小册子《论概率博弈的计算》,该书出版于 1657 年.

惠更斯的著作仅包含 14 个命题并以 5 个给读者的练习结束.命题包含有对德·默勒两个问题的处理,但惠更斯还给出了对解法背后的推理的详细讨论,特别是计算概率博弈的方法:"虽然在一个纯概率博弈中结果是不确定的,但一个参与者赢或者输的机会取决于一个确定的值."[34] 惠更斯的"值" 同帕斯卡押宝的概念类似,但在概率博弈中,惠更斯可以明确地计算它.用现代的术语说,一个机会的"值" 就是**期望**,即一个人如果进行许多次某一赌博可以赢取的平均数量.一个人大概愿意支付这一数量以取得玩一个公平赌博的机会.例如,惠更斯的第一个命题是:"能以相等的机会赢取 a 或者 b 的量对我的价值是 $(a+b)/2$."[35] 这一命题同帕斯卡解决分配问题时陈述的一条原理相同.但惠更斯给出了一个证明.他假定两个人每人出 $(a+b)/2$ 的赌注,两人获胜的机会均等.如果第一人赢,他得到 a 而他的对手得到 b.如果第二人赢,支付则相反.惠更斯认为这是一个公平博弈.使用现代的术语,因为赢取 a 或者 b 的概率是 1/2,每个参与人的期望都是 $(1/2)a+(1/2)b$,即惠更斯的机会的"值".

惠更斯在他的第三个命题中对该结果作了推广:"有 p 次机会赢得 a 以及 q 次机会赢得 b,机会都是同样的,对我的价值是 $(pa+qb)/(p+q)$."[36] 换句话说,如果 $p+q=r$,如果赢得 a 的概率是 p/r,赢得 b 的概率是 q/r,则期望是 $(p/r)a+(q/r)b$.通过把这一问题嵌入到一个由排成一个圆圈的 $p+q$ 个人参与的对称博弈的方法证明了这一结果,每个参与人投入同样的赌注 x 并且每人获胜的可能性相等.[37] 如果一个确定的参与人获胜,他拿走全部赌注,给他左边的 $q-1$ 参与人每人支付 b,给他右边的 p 个人每人支付 a,并保留其余.为使得这一余额等于 b,必须成立

$$(p+q)x-(q-1)b-pa=b \quad \text{或者} \quad x=\frac{pa+qb}{p+q}.$$

但很清楚每个参与者现在都有 q 个机会赢取 b 以及 p 个机会赢取 a,所以博弈是公平的,并且每个参与人应该愿意以提到的赌注冒险.

惠更斯把每个公平博弈的参与人愿意拿出经过计算的公平的赌注冒险而不愿意拿出更多的数量作为一个公理.但事实上,正如赌博的历史所显示,这一假定至少也是有争议的.惠更斯定义的公平赌注是一个人为参加赌博的机会愿意支付的最多的数额压根并不清楚.国家经营的彩票的成功证实了恰恰相反的东西,更不用说拉斯维加斯和大西洋城的大赌场了.无论如何,惠更斯将他论文的剩余部分建立在第三个命题的基础之上,并且期望的概念即使在今天也被认为是一个有用的概念.

惠更斯对德·默勒分配问题的讨论同帕斯卡的相似,但他在命题 11 中对问题给出了更为广泛的分析.他说明了如何确定两枚骰子被掷的次数,使得一个人为在这么多次投掷中出现两个 6 点时可以赢到 a 而愿意出 $(1/2)a$.惠更斯分步进行讨论.假定当两个 6 点出现时一个人可以赢到 a,他推断说在第一次投掷时,一个人有 1 次机会赢取 a,而有 35 次机会赢取 0,所以 1 次投掷机会的值为 $(1/36)a$.如果参与人第一掷失败,他进行第二掷,这一掷的值自然也同样是 $(1/36)a$.所以对第一掷来说,参与人有一次机会赢取 a 以及 35 次进行值为 $(1/36)a$ 的第二次投掷.根据第三命题,在两次投掷中掷出 2 个 6 点的机会的值为:

$$\frac{1a+35(1/36)a}{1+35} \quad \text{或者} \quad (71/1296)a.$$

惠更斯然后过渡到 4 次投掷的情况.如果一个参与人在头两次投掷中掷出两个 6 点,他赢到 a;如果掷不出,他得到第二次值为 $(71/1296)a$ 的两次投掷的机会.因为在头两次投掷中赢取 a 有 71 次机会因此赢不到的机会是 1225 次(在 1296 次当中),进入第二轮的机会是 1225 次,它们的值也是

(71/1296) a. 还是根据第三命题，参与人在 4 次投掷中出现两个 6 点的机会的值是：

$$\frac{71a + 1225(71/1296)a}{1296} \quad 或者 \quad \frac{178\,991}{1\,679\,616}a.$$

因为这个值仍然显著地小于希望的 (1/2) a，惠更斯不得不继续这一过程. 虽然他没有给出进一步的计算，他写到人们然后考虑 8 次，然后 16 次，再然后是 24 和 25 次投掷. 结果表明在 24 次投掷中参与人赌 (1/2) a 稍稍不利而在 25 次投掷中又占些许便宜.[38]

在他短小精悍的论文末尾，惠更斯提出了若干从瓮中抽取带色小球的问题作为练习，这一类型的问题今天在每种初等概率论的课本中出现. 这些问题在随后的几十年里被许多数学家讨论，特别是因为惠更斯的课本直到 18 世纪早期仍是能找到的概率论的唯一入门教程. 即使到了 18 世纪早期之后，它的影响还在延续，因为詹姆斯 · 伯努利把它吸收进他自己关于概率论的更为广泛的著作中，这就是 1713 年的《猜度术》.

11.4 数 论

参与过解析几何和概率论始创的费马对数论也作出了贡献，这些贡献在他生前实际上直到下个世纪中期基本上被忽略了. 忽略的原因之一也许是他对他的方法讳莫如深. 因此，尽管他的许多结果为人所知，因为他在给各个通信伙伴的信中骄傲地宣布过这些结果并同时提出类似的题目作为挑战，实际上没有他的任何证明的记录，只有某些方法的模糊轮廓.

费马对数论的最早兴趣源自完全数的经典概念，即那个等于它所有真因子之和的数. 欧几里得《几何原本》的第九卷中包含了，如果 $2^n - 1$ 是素数，则 $2^{n-1}(2^n - 1)$ 为完全数的一个证明. 但希腊人只能发现 4 个完全数，6，28，496 和 8128，因为很难确定使得 $2^n - 1$ 为素的 n 的值. 费马发现了可以在这方面有所帮助的 3 个命题，他在 1640 年 6 月给梅森的一封信中传达过这些命题. 这第一个结果是，如果 n 本身非素，则 $2^n - 1$ 不可能为素. 这一结果的证明只需要给出它的因子：如果 $n = rs$，则

$$2^n - 1 = 2^{rs} - 1 = (2^r - 1)(2^{r(s-1)} + 2^{r(s-2)} + \cdots + 2^r + 1).$$

基本问题因此归结为对什么样的素数 p，$2^p - 1$ 是素数. 为纪念费马最喜爱的通信伙伴，这样的素数今天被称为梅森素数.

费马的第二个命题是说，如果 p 是一个奇素数，则 $2p$ 整除 $2^p - 2$ 或者说 p 整除 $2^{p-1} - 1$. 他的第三个命题是，在同样的假设下，$2^p - 1$ 唯一可能的因子具有形式 $2pk + 1$. 费马在信中没有暗示这些结果的任何证明而仅仅给出了若干数值例子. 通过用形为 $74k + 1$ 的数检验整除性而最终发现了因子 $223 = 74 \cdot 3 + 1$，他证实 $2^{37} - 1$ 为一合数. 但在数月之后写给贝尔纳 · 弗里尼科 · 德 · 拜西 (1612—1675) 的一封信中，他叙述了一个更为普遍的定理，以上两个命题是该定理的简单的推论. 这一今天以费马小定理闻名的定理用现代的术语表述就是，如果 p 是任一素数而 a 是任一正整数，则 p 整除 $a^p - a$. (这一定理经常被写作 $a^p \equiv a (\mathrm{mod}\ p)$ 的形式，或者加上 a 与 p 互素的条件，写成 $a^{p-1} \equiv 1 (\mathrm{mod}\ p)$. 可以推出，如果 n 是使得 p 整除 $a^n - 1$ 的最小正整数，则 n 整除 $p - 1$，并且所有满足 p 整除 $a^k - 1$ 的指数 k 都是 n 的倍数.) 费马没有在他的任何书面材料中给出他如何发现或者证明这一结果的任何暗示. 无论如何，在给梅森信中的第二个命题不过是定理在 $a = 2$ 时的情形 (这里 $p > 2$). 第三个命题只需要多一点的工作. 假定 q 是 $2^p - 1$ 的一个素因子，那么定理意味着 p 整除 $q - 1$ 或者 $q - 1 = hp$，h 是某一个整数. 因为 $q - 1$ 是偶数，2 必定整除 hp 因而整除 h. 由此推出命题断言的 $h = 2k$ 或者 $q = 2kp + 1$.

结果表明费马小定理是数论中一个有着许多应用的极端重要的结果. 但费马关于素性另一方

面的工作却说明即使是费马也是会犯错误的.费马在他的通信中一再声称所谓的费马数,就是形式为 $2^{2^n}+1$ 的数,都是素数.直到 1659 年,他还写道他发现了一个证明.不难说明这种形式的数在 $n=0,1,2,3,4$ 时是素数.但莱昂哈德·欧拉在 1732 年发现 641 是 $2^{2^5}+1$ 的一个因子,并且实际上还没有发现大于 $2^{2^4}+1$ 的费马数.费马怎么会犯这样的错误?很可能是他尝试的证明同他在数论工作的另一领域勾画的方法属于同一种类型,无穷递降法,并且他轻易地相信对 4 以下整数适用的方法对更大的整数也适用.

无穷递降法包含在费马实际详细写出的唯一一个数论证明中,问题是找出一个面积为给定数值的整数直角三角形.费马评论说这样一个三角形的面积不可能为平方数,就是说,不可能找到整数 x,y,z,w 使得 $x^2+y^2=z^2$ 且 $(1/2)xy=w^2$.费马用来证明这一结果的无穷递降法证明不存在具有某一性质的正整数的方法是:说明一个整数具有这一性质的假设意味着一个更小的整数也具有同样的性质.继续这一论证,人们就可以得到一个无限递减的正整数列,这是不可能的.

在这一具体情况下,费马首先说明,如果某一个由互素且奇偶性相反的数对 p,q 产生的毕达哥拉斯三元数满足提到的条件,则存在两个相差一个平方数的 4 次方数.理由是,因为 $x=2pq$ 且 $y=p^2-q^2$,面积 $(1/2)xy=pq(p^2-q^2)$ 将是一个平方数.所以,$p=d^2,q=f^2$,并且依条件 $p^2-q^2=d^4-f^4=c^2$.然后,费马注意到,因为 $c^2=(d^2+f^2)(d^2-f^2)$ 并且因为 d 与 f 互素,d^2+f^2 与 d^2-f^2 必定都是平方数,令 $d^2+f^2=g^2$ 且 $d^2-f^2=h^2$.从第一个方程中减去第二个得 $2f^2=g^2-h^2=(g+h)(g-h)$.因为 g^2 与 h^2 都是奇数并且互素,$g+h$ 与 $g-h$ 都是偶数并且没有 2 以外的公因子.由此推出 $g+h$ 可以写成 $2m^2$,而 $g-h$ 可写成 n^2(或者相反),n 是偶数而 m 是奇数.所以 $g=m^2+n^2/2,h=m^2-n^2/2$ 以及 $d^2=(1/2)(g^2+h^2)=(m^2)^2+(n^2/2)^2$,但这时 m^2 和 $n^2/2$ 成为面积 $m^2n^2/4$ 也是平方数的新的直角三角形的两边.因为这一新的三角形的斜边 d 小于原三角形的斜边,无穷递降法推出原假定必定为假.

使用无穷递降法人们可以从这一论证中引出一个论证,说明无法找到三个正整数 a,b,c 使得 $a^4-b^4=c^2$.由此推出人们也无法将一个四次方表达为另两个四次方之和.这一结果的推广,推广到"人们不能把立方数拆成两个立方数,也不能把一个四次方数拆成两个四次方数,一般地,不能把超过平方直到无穷的任何次方数拆成两个同次方数之和,"[39] 费马把这段话作为对丢番图问题Ⅱ-8的一个边注写在他 1621 年的拉丁版《算术》上,这构成了以费马大定理闻名的定理的内容(图 5.4).用现代的术语,这一猜想断言不存在非零整数 a,b,c 以及 $n>2$ 使得 $a^n+b^n=c^n$.对这一结果,费马声称他有"一个真正美妙的证明 …… 这里的空白太狭小而无法容纳," 这给 17 世纪以来的数学家们提出了重要的挑战.1995 年,普林斯顿大学的安德鲁·怀尔斯(生于 1953 年)给出了费马大定理的第一个证明,该证明建立在 20 世纪后期的许多其他数学家的工作之上并用到了费马所没有的代数几何的技巧.因此大多数历史学家相信费马在声称自己有了证明时弄错了,可能是因为他错误地以为适用于 $n=3$ 和 $n=4$ 情形的无穷递降法可以推广到更大的 n 值.

虽然费马对费马数的声言是错误的并且他断言费马大定理成立是不成熟的,但他在通信中或者涂写在丢番图著作的边角处的对绝大多数数论结果的声言被证明是成立的.虽然他试图在许多场合唤起其他欧洲数学家来研究他提出的各种数论问题,但他的呼吁却得不到任何响应.直到下一个世纪,由这个法国律师开创的数论工作才找到了后继者.

11.5 射影几何

被忽视的命运也降临在了吉拉德·德扎格(Girard Desargues,1591—1661)这位法国工程师和建

筑师的头上,德扎格对数学最有创造性的贡献是在射影几何领域.作为他职业兴趣的一部分,他希望能继续由文艺复兴艺术家开创的对射影的研究.在掌握了希腊人的几何工作特别是阿波罗尼乌斯的工作后,他打算统一各种方法,不是像费马那样将其代数化而是将它们包含在新的综合的射影技巧下.特别是,他试图在他 1639 年的《处理圆锥同平面相交结果的草案》中用射影技巧统一对圆锥曲线的研究.例如,众所周知,圆从斜的方向看像个椭圆.因为从斜的方向看等价于从圆以外的平面上的一点将圆投影到另一平面,德扎格希望研究这些在射影条件下保持不变的圆锥曲线的性质.

作为他研究的一部分,德扎格不得不考虑无穷远处的点,它们就像透视图中的灭点,是平行线相交的地方."如果需要,任一直线都可以看作在两个方向上通向无穷." 当几条直线平行或者相交于同一点时,德扎格说它们属于同一**布局**."因此同一平面中的任意两条直线属于同一布局,它们的交点在有限远处或者无限远处."[40] 所有无限远处点的集合构成无限远直线.由此推出所有的平面必须被理解成在所有方向上伸展到无穷远.另外,因为圆柱可以被看作是顶点在无穷远的圆锥,德扎格同时处理圆锥和圆柱.因此,两条圆锥截曲线通过从顶点的射影变换相联系,而两条圆柱截曲线通过从无限原点的射影相联系.因为圆是一种圆锥(或圆柱)截曲线,德扎格能够把任何圆锥曲线看作是同圆射影等价.椭圆是在它们的平面上碰不到无穷远直线的射影,抛物线是仅仅碰到无穷远直线的射影,而双曲线是同无穷远直线相交的那些圆的射影.因此,圆的任何在射影下不变的性质都可以轻易地被证明是所有圆锥曲线的性质.

但是德扎格最著名的结果没有出现在《草案》中,而是出现在一个叫做亚伯拉罕·波色(1602—1676)的朋友所写的一本实用著作《德扎格应用透视的普遍方法》的附录中:"当在相同或不同平面的直线 HDa, HEb, cED, lga, lfb, HlK, DgK, EfK,不论它们的顺序或者方向如何,相交于相似的点,则点 c, f, g 位于同一直线 cfg 上."[41] 用现代的术语,德扎格考虑了两个三角形 KED 和 abl,它们通过从"相似的"点 H 的射影相联系(图 11.16).换句话说,连接各对应顶点的直线交于 H.结论是这时对应边——这里是 DK, al; EK, bl;与 DE, ab——的交点 g, f, c 都位于同一直线上.德扎格运用梅内劳斯定理证明了这一结果.

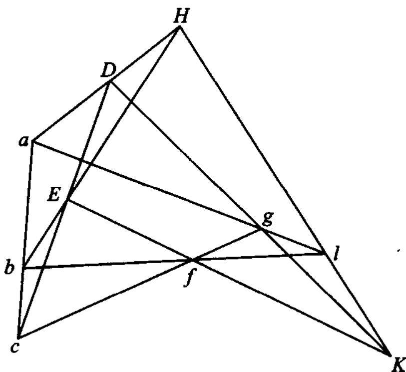


图 11.16 德扎格定理.

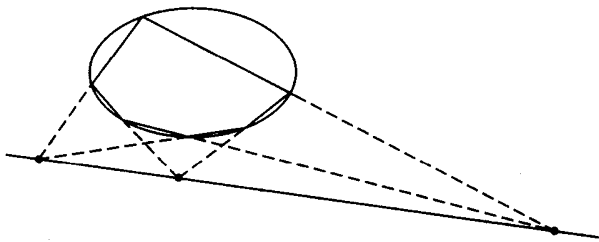


图 11.17 帕斯卡的六边形定理.

德扎格的工作没有被得到很好的承认,这一方面是因为他发明并使用了如此之多的新的技术术语以至于很少有人能理解它;另一方面是因为数学家们才刚开始领悟到笛卡儿用解析方法统一几何的奥妙而不准备考虑一种新的综合的统一方式.很明显欣赏他的工作的唯一一个同时代的数学家是帕斯卡,帕斯卡在 1640 年发表了一篇题为《圆锥曲线论》的短文,在其中他感谢德扎格使他

了解了射影方法.这一著作包含了此后冠以帕斯卡名字的定理.

定理　如果一个六边形内接于一个圆锥曲线,则它的对边相交于三个共线点(图 11.17).

因为该定理是作为射影几何的命题提出的,所以可能的情形中包括某些对边平行因而交点在无穷远处的情况.这自然是六边形是圆内接正多边形的情况.帕斯卡在他的短文中没有给出定理的证明.他仅仅先就圆的情形而后就任意圆锥曲线的情形声称定理成立.据推测,他打算沿着德扎格设想的方案证明这个普遍的结果.帕斯卡许诺在一本关于圆锥曲线的更完整的著作中透露更多的结果和方法,他在 17 世纪 50 年代中期完成了这一著作.不幸的是,这一较为详尽的著作从未出版,而且所有的手稿后来都不见了.事实上,几何中的射影方法在 19 世纪早期之前实际是被忽略了.

习　题

来自费马《概论》的问题

1.假定 $xy = c$ 表示一条渐近线为 x - 轴和 y - 轴的双曲线,证明 $xy + c = rx + sy$ 也表示双曲线.找出它的渐近线.

2.确定方程 $b^2 - 2x^2 = 2xy + y^2$ 的轨迹.(提示:给两边加 x^2.)

3.证明 $b^2 + x^2 = ay$ 表示抛物线.画出它在第一象限的部分.

4.确定解决了两点情况下的阿波罗尼乌斯的问题的圆的方程.(令图 11.1 中 A 和 B 点的坐标分别是和 $(-a,0)$ $(a,0)$).

5.确定解决了四个非共线点 $(x_i, y_i)(i = 1,2,3,4)$ 情况下的阿波罗尼乌斯的《平面轨迹》中问题的圆的方程.

来自笛卡儿《几何学》的问题

6.使用原著中的各常量,确定在前两条线段的积等于后两条线段的积的特殊情况下解决了 4 线问题的轨迹的方程.这是什么曲线?

7.证明图 11.6 中的 MQ 和 MR 表示方程 $z^2 = ax - b^2$ 的两个解.

8.怎样利用图 11.6 表示 $z^2 = ax + b^2$ 的负解?

9.为解 4 次方程 $x^4 - px^2 - qx - r = 0$,笛卡儿考虑 y^2 的 3 次方程:$y^6 - 2py^4 + (p^2 + 4r)y^2 - q^2 = 0$.假如 y 是一个解,证明原多项式可以分解为两个二次多项式:$r_1(x) = x^2 - yx + 0.5y^2 - 0.5p - 0.5q/y$,$r_2(x) = x^2 + yx + 0.5y^2 - 0.5p + 0.5q/y$,任何一个都可解.应用这一方法解方程 $x^4 - 17x^2 - 20x - 6 = 0$.注意对应的 y 的方程 $y^6 - 34y^4 + 313y^2 - 400 = 0$ 有解 $y^2 = 16$.

10.解方程 $x^3 - \sqrt{3}x^2 + \frac{26}{27}x - \frac{8}{27\sqrt{3}} = 0$,首先作代换 $y = \sqrt{3}x$ 然后作代换 $z = 3y$ 得到一个关于 z 的整系数方程.

11.用笛卡儿的符号法则研究 $2x^4 - 9x^2 - 5x + 1$ 的根的性质.

12.作出笛卡儿对三次方程 $x^3 = -4x + 16$ 的图解.作出海亚姆对同一个方程的图解.哪种方法更简单?

13.阅读笛卡儿《几何学》第三编中详细给出解 4 次多项式方程的图解法的部分.概述该方法并用它解 $x^4 = x^2 + 5x + 2$.

德·威特的问题

14.在德·威特化简方程 $y^2 + \frac{2bxy}{a} + 2cy = bx - \frac{b^2x^2}{a^2} - c^2$ 所作的代换 $z = y + bx/a + c$ 中,他把一个坐标轴旋转了一个角度 α,求该角的正弦和余弦.

15. 证明德·威特的方程 $y^2 + \frac{2bxy}{a} + 2cy = \frac{fx^2}{a} + ex + d$ 表示一条双曲线.(利用代换 $z = y + bx/a + c$ 并说明这一代换同形如 $x' = \beta x$ 的代换相结合,可以将原来的斜的 $x - y$ 坐标系转化为新的以垂直坐标轴为基础的新的 $x' - z$ 坐标系.)画出曲线的草图.

吉拉德的问题

16. 已知 $x = 18$ 是一个解,用吉拉德的技巧解方程 $x^3 = 300x + 432$.
17. 用吉拉德的技巧解方程 $x^3 = 6x^2 - 9x + 4$,首先用观察的方法确定一个解.
18. 证明在方程 $x^4 + Bx^2 + D = Ax^3 + Cx$ 中,A 是根的和,$A^2 - 2B$ 是根的平方和,$A^3 - 3AB + 3C$ 是根的立方和而 $A^4 - 4A^2B + 4AC + 2B^2 - 4D$ 是根的 4 次方的和.
19. 这一问题说明吉拉德对多项式方程负数解的几何解释.设两直线 DG, BC 在点 O 处相交成直角(图 11.18).在平分直角于点 O 的直线上确定点 A 使得 $ABOF$ 是边为 4 的正方形.如图中作 ANC 使得 $NC = \sqrt{153}$.求 FN 的长.(吉拉德注意到如果 $x = FN$,则 $x^4 = 8x^3 + 121x^2 + 128x - 256$ 因此存在 4 个可能解,每一个都能被计算.两个正的可能解被表示为 FN 和 FD,而两个负解被表示为 FG 和 FH,后两个的取的方向同前两个相反.)

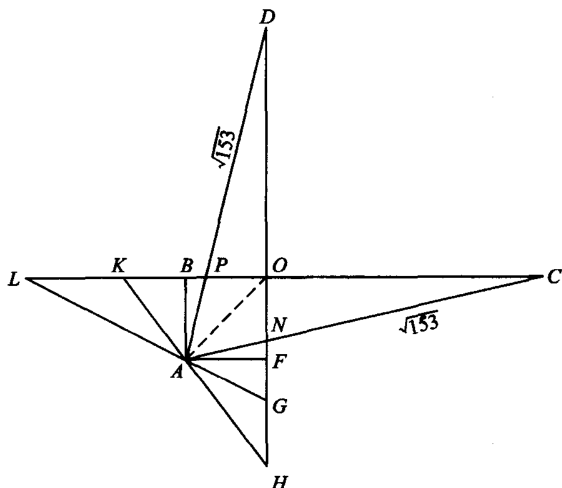


图 11.18　吉拉德的问题.

帕斯卡的问题

20. 对 n 用归纳法证明对于所有小于 n 的 k,$\binom{n}{k} = \sum_{j=k-1}^{n-1} \binom{j}{k-1}$.
21. 证明 $\binom{n}{k} : \binom{n}{k+1} = (k+1) : (n-k)$.
22. 证明 $\binom{n}{k} : \binom{n-1}{k} = n : (n-k)$.
23. 证明对于所有小于 n 的 j,$\sum_{k=0}^{j} \binom{n}{k} = \sum_{k=0}^{j} \binom{n-1}{k} + \sum_{k=0}^{j-1} \binom{n-1}{k}$.
24. 帕斯卡断言投掷一个骰子 4 次出现一个 6 的胜负比是 671 : 625.说明这为什么成立.
25. 说明投掷 3 枚骰子一次至少出现一个 1 点的负胜比是 125 : 91.(这一答案是卡尔达诺给出的.)
26. 确定当以下两人赌局中断时合理分配赌注的方案:甲需要再胜 3 局而乙需要再胜 4 局.
27. 假定 3 人在玩一组公平博戏,规定首先赢 3 场的人得到赌注.如果他们中断游戏时,甲需要再赢 1 场,而乙和丙均需再赢两场;找到合理分配赌注的方案.(帕斯卡和费马的通信中曾经讨论过该问题.)

28. 投掷 3 枚骰子时,取得 9 和 10 各有 6 种不同方式.然而,说明掷出 10 的概率高于掷出 9 的概率.(伽利略著作的一个残篇中发现有对该思想的讨论.)
29. 对过圆锥曲线上一点 P 作曲线的切线用帕斯卡的 6 边形定理.将切线考虑为通过 P 的两个邻点.然后在圆锥曲线上取另外 4 个点并应用帕斯卡的定理.

惠更斯的问题

30. 两人玩掷双骰游戏,骰子的点数和为 7 时甲赢,为 6 时乙赢,为其它时两人平分赌注,求两人的期望(机会的价值).
31. 假如我同别人玩轮流掷双骰的游戏,条件是如果我掷出 7 点则我胜,他掷出 6 点则他胜,如果他先掷,我的机会同他的机会的比是多少?
32. 在一瓮中有 12 个球,4 白 8 黑.3 个蒙住了眼睛的游戏参加人轮流取球,先甲、再乙、后丙.首先取出白球的人获胜.假定黑球取出后都会被放回,求三人取胜的机会的比.
33. 有 40 张牌,每种花色 10 张.甲同乙打赌他能抽出 4 张牌,每种花色 1 张.每人投入的赌注应该是多少?

费马数论中的问题

34. 通过将 $2^p \equiv (1+1)^p$ 用二项式定理展开并注意到所有满足 $1 \leqslant k \leqslant p-1$ 的二项式系数 $\binom{p}{k}$ 都能被 p 整除,证明如果 p 是素数,则 $2^p \equiv 2 \pmod p$.用这一结果和 $(a+1)^p \equiv a^p + 1 \pmod p$ 的事实,通过对 a 归纳证明 $a^p \equiv a \pmod p$.
35. 作为当 a 和 p 互素情形下费马小定理的证明,考虑用 p 除 $1, a, a^2, \cdots$ 所得的余数.这些余数最终必定重复(为什么?),并且因此 $a^{n+r} \equiv a^r \pmod p$ 或者 $a^r(a^n - 1) \equiv 0 \pmod p$ 或者 $a^n \equiv 1 \pmod p$.(说明每种可能的合理性.)取 n 作为满足最后一个同余式的最小正整数.应用辗转相除算法,说明 n 整除 $p-1$.

讨　论

36. 笛卡儿《方法论》中最有名的一段引言是:“我思故我在”.它的上下文是笛卡儿决心只接受那些自明的观念为真.以这段引言为基础有一个有名的幽默:笛卡儿走进一家饭店.侍者问他:“您想要点今晚的特色菜吗?”他回答说:“我不想”,接着就消失了.评价这一幽默逻辑上的合理性.[42]
37. 比较笛卡儿、费马和德·威特的解析几何.将其中一个作者的表述加以改造用来给尚未学过微积分的学生作关于这一学科的一个讲座.
38. 哪些技术以及/或者理解上的进展使得吉拉德能够表述出任何一个代数方程都有同它次数一样多的解?为什么卡尔达诺和韦达不能做到这点?
39. 规划一堂方程理论课,用笛卡儿的代数技巧讲授诸如因子定理以及解高于二次的多项式方程的方法之类的结果.
40. 用卡尔达诺的思想规划一堂基础概率论课.将说明涉及到的各种规则的合理性以及人们可能会犯的错误的材料包括进去.
41. 使用帕斯卡《论算术三角形》中的材料规划一堂关于数学归纳法原理的课.
42. 将帕斯卡对数学归纳法的运用同哈塞姆(ibn al - Haytham),塞毛艾勒以及莱维对归纳法的运用加以比较.(关于欧洲归纳法起源的更多的信息,见 W.H.Bussey, “The Origin of Mathematical Induction”, *American Mathematical Monthly* 24(1917), 199 - 207.)
43. 找出谁最先用“数学归纳法”这一术语描述帕斯卡和其他人使用的技巧.为什么选择这一术语?(见 Florian Cajori, “The Origin of the Name'Mathematical Induction',” *American Mathematical Monthly* 25(1918), 197 - 201.)

文献和注解

解析几何的唯一一部通史是 Carl Boyer 的 History of Analytic Geometry(New York:Scripta Mathematica, 1956).这部著作的范围覆盖了该学科从古代的开端一直到 19 世纪的发展.从 The Scholar's Bookshelf 可以得到该书新近的一个重印本.对费马和笛卡儿工作的一个简要考察包含在 Carl Boyer 的"Analytic Geometry:The Discovery of Fermat and Descartes", Mathematical Teacher 37(1944),99 - 105 和"The Invention of Analytic Geometry", Scientific American 180(January 1949),40 - 45.对 17 世纪以及更早的时期有所论述的最好的一本概率论的通史是 F.N.David 的"Games, Gods, and Gambling"(New York: Hafner,1962) 和 Ian Hacking 的"The Emergence of Probability"(Cambridge:Cambridge University Press,1975).后一本书更富哲学气息,而前者则更详尽地讨论相关著作.用历史的观点写成的一部数论通史是 André Weil "Number Theory:An approach through history from Hammurapi to Legendre"(Boston:Birkhäuser,1983).对德扎格和帕斯卡在射影几何上工作的讨论见 J.V.Field 和 J.J.Gray,"The Geometrical Work of Girard Desargues"(New York:Springer,1987).

1. David Eugene Smith, *A Source Book in Mathematics*(New York:Dover,1959),vol.2,p.389.
2. 对费马的数学工作和生平最好的研究见 Michael S.Mahoney, *The Mathematical Career of Pierre de Fermat 1601—1665*(Princeton:Princeton University Press,1973).该书包含有对费马工作的详尽的分析,不仅有在解析几何方面的工作而且还有将在第 12 章中论述的在微积分各方面的工作.
3. 同上,102 - 103.
4. Smith,Source Book,p.390.
5. 同上,p.392.
6. Mahoney,The *Mathematical Career*,p.91.
7. René Descarte, *Discourse on Method*, *Optics*, *Geometry and Meteorology* translated by Paul J.Olscamp(Indianapo-lis:Bobbs-Merrill,1965),p.9.
8. David Eugene Smith 和 Marcia L.Latham 翻译的 *The Geometry of René Descartes*(New York:Dover,1954),p.2 这是《几何学》的一标准的现代版本.它包括了原著扉页上的法语和英文翻译以及笛卡儿原始的注释的插图.
9. Smith 和 Latham, *Geometry*,p.5.
10. 同上,p.22.
11. 参见 H.J.M.Bos 非常有趣的论文"On the Representation of Curves in Descarts' Géométrie" *Archive for History of Exact Sciences* 24(1981),295 - 338.Bos 讨论了笛卡儿在《几何学》中勾画出的对几何学的一般纲领.也见 A.Molland 的"Shifting the Foundations:Descartes'Transformation of Ancient Geometry" *Historia Mathematica* 3(1976),21 - 49.见 E.G.Forbes 的"Descates and the Birth of Analytic Geometry" *Historia Mathematica* 4(1977),141 - 161,该文提出解析几何应该归功于韦达的一个学生 Marino Ghetaldi(1566 - 1627).
12. Smith 和 Latham, *Geometry*,p.43.想找有关笛卡儿的曲线绘制仪器的更多的信息,见 David Dennis 的"René Descartes' Curve-Drawing Devices:Experiments in the Relations Between Mechanical Motion and Symbolic Language." *Mathematics Magazine* 70(1997),163 - 175.
13. 同上,p.48.
14. 同上,p.91.
15. 同上,p.240.
16. 想要哈略特更多的资料,参见 J.A.Lohne 的"Dokumente zur Revalidierung von Thomas Harriot als Algbraiker" *Archive for History of the Exact Sciences* 3(1996),185 - 205 以及"Essays on Thomas Harriot" *Archive for History of the Exact Sciences* 20(1979),189 - 312.
17. *The Early Theory of Equations*: *On their Nature and Constitution*(Annapolis:Golden Hind Press,1986),p.107.该书包括

爱仑 · 布莱克翻译的吉拉德的 A New Discovery in Algebra 以及韦达和德邦(Debeaune) 的论文.

18.同上,p.145.

19.同上,p.138.

20.同上,p.139.

21.同上,p.141.

22.Smith 和 Latham, Geometry, p.159.

23.同上,p.175.

24.同上,p.159.

25.同上,p.160.

26.同上,p.200.

27.同上,p.240.

28.对犹太人对概率的早期贡献的讨论见 Nachum L. Rabinovitch, *Probability and Statistical Inference in Ancient and Medieval Jewish Literature*(Toronto: University of Toronto Press, 1973).

29.转引自 David 的 *Games*, *Gods*, p.33.

30.Oystein Ore, *Cardano*, *the Gambling Scholar*(Princeton: Princeton Universtity Pres, 1973), p.203.这部传记著作的末尾附有 Sydney Henry Gould 翻译的 Gerolamo Cardano 的《论机会游戏》.Ore 详细讨论了卡尔诺的工作并比 David 和 Hacking 更强烈的主张是卡尔达诺创造了概率论的许多核心思想.

31.转引自 Oystein Ore 的"Pascal and the Invention of Probability Theory" *American Mathematical Monthly* 67(1960), 409 – 419, p.414.

32.布莱西 · 帕斯卡的《论算术三角形》, Richard Scofield 翻译,见 *Great Books of the Western World*(Chicago: Encyclopedia Britannica, 1952), vol.33, p.452.该版本还包含有帕斯卡和费马关于概率论的通信.更多帕斯卡的资料见 Harold Bacon 的"The Young Pascal" *Mathematical Teacher* 30(1937), 180 – 185. Morris Bishop 的 *Pascal*, *The Life of Genius*(New York: Reynal and Hitchcock, 1936) 以及 Jean Mesnard 的 Pascal, His Life and Works(New York: Philosophical Library, 1952).

33.同上.

34.Christian Huygens, *On the Calculations in Games of Chance in Oeuvres completes*(The Hague: 1888 – 1950), vol.14, p.61.

35.同上,p.62.

36.同上,p.64.

37.惠更斯版中关于他的博弈是不完整的.这一改进见 Olav Reiersol 的"Notes on some Propositions of Huygens in the Calculus of Probability", *Nordisk Matematisk Tidskrift* 16(1968), 88 – 91.

38.对德 · 默勒骰子问题的现代讨论见 Jane B. Pomeranz"The Dice Problem-Then and Now", *College Mathematics Journal* 15(1984), 229 – 237.

39.转引自 Mahoney 的 *Mathematical Career*, p.344.关于费骊无穷递降法的更多的信息见 Howard Eves 的"Fermat's Method of Infinite Descent" *Mathematics Teacher* 53(1960), 195 – 196.另见 Michael Mahoney 的"Fermaet's Mathematics: Proofs and Conjectures" *Science* 178(1972), 30 – 36.

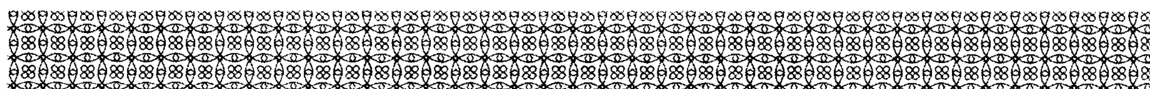
40.Field 和 Gray 的 *Geometrical Work*, pp.69 – 70.有关德扎格的更多信息见 N.A.Court 的"Desargues and his Strange Theorem", *Scripta Mathematica* 20(1954), 5 – 13, 155 – 164.

41.Smith, *Souree Book*, p.307 – 308.

42.我的儿子阿里讲给我的.

17 世纪的几何、代数和概率论概览

1501—1576	卡尔达诺(Girolamo Cardano)	概率论
1560—1621	托马斯·哈略特(Thomas Harriot)	方程理论
1588—1648	玛林·梅森(Marin Mersenne)	活的科学杂志
1591—1661	吉拉德·德扎格(Girard Desargues)	射影几何
1595—1632	阿尔伯特·吉拉德(Albert Girard)	方程理论
1596—1650	笛卡儿(René Descartes)	解析几何、方程理论
1601—1665	费马(Pierre de Fermat)	方程理论、概率论、数论
1607—1684	安妥尼·哥保德(Antoine Gombaud)	概率论
1615—1660	范·舒滕(Frans van Schooten)	解析几何
1623—1662	帕斯卡(Blaise Pascal)	概率论、射影几何
1623—1672	德·威特(Jan de Witt)	解析几何
1629—1695	惠更斯(Christian Huygens)	概率论



第12章 微积分的开端

> 我的创造性全盛时期来到时，我学习几何还不到6个月，在此期间我发现了特征三角形方法和另外一些诸如此类的事项…… 当时，我对笛卡儿代数还相当无知，也不了解不可分量方法；实际上我不知道重心的正确定义.因为，当我偶然地向惠更斯说起它时，他发觉我以为通过重心的直线总是将图形分为面积相等的两部分…… 惠更斯听到这点大笑起来，他告诉我再没有什么事情比这更荒谬的了.于是我在这一事情的强烈激励下开始致力于更为精深的几何的研究.
>
> ——选自1680年戈特弗里德·莱布尼茨给埃仑弗里德·瓦尔特·冯·切恩豪斯(1651—1708)的一封信[1]

牛顿在1676年10月24日通过亨利·奥尔登堡(Henry Oldenburg,1615—1677)发给莱布尼茨的第二封信(也是最后一封，称为“后信”)中用密码隐喻了他的微积分的基本目标，以防泄露太多的秘密.该密码记为6accdæ13eff7i319n4o44qrr4s8t12ux，很难相信莱布尼茨能把它解读为“给定一个含有任意多个流量的方程，求流数，以及相反的问题”.[2] 尽管莱布尼茨就像回复前信一样热情地回复了牛顿的信，给出了他自己在微积分方面工作的细节并且邀请他进一步对话，牛顿始终没有回复.

几个世纪以来，许多数学家考虑过确定由曲线围成区域的面积以及寻找某些函数的最大或者最小值的问题，在他们工作的基础上，17世纪后半叶的两个天才人物，伊萨克·牛顿和戈特弗里德·莱布尼茨，创造了微积分方法，该方法是现代数学分析的基础以及在数目不断增加的其它学科应用的源泉.最大－最小值问题、面积问题以及相关的寻找切线和确定面积的问题许多年来就被探索过并在各种特殊情形下获得解决.但是在几乎所有被解决的例子中，不论是在前希腊时期、由希腊人自己还是由他们的伊斯兰后继者，解法都需要天才的构思.没有人提出过一种能使这些问题在新情况下被轻易解决的算法.

不论在希腊还是在伊斯兰的背景下，新情况都不会经常出现，因为这些数学家们几乎没有什么办法描述要被计算切线、面积或者体积的新的曲线或者立体.但随着解析几何在17世纪上半叶的

发明,任何种类的新曲线和立体的构造一下子成为可能.无论如何,任何代数方程都确定一个曲线,而新的立体可以被构造,比方说,通过围绕同一平面任何直线旋转曲线的方法.有了无数多个供研究的新的例子,17 世纪的数学家们寻找并且发现了求最大值、作切线以及计算面积和体积的新的方法.但是,这些数学家不关心函数.他们关心由两个变量间的某些关系定义的曲线.在寻找切线的过程中,他们经常考虑曲线的其它几何层面.图 12.1 描绘了同给定曲线上的一个点相联系的一些量:横坐标 x,纵坐标 y,弧长 s,次切距 t,切线 τ,法线 n 以及次法距 v.

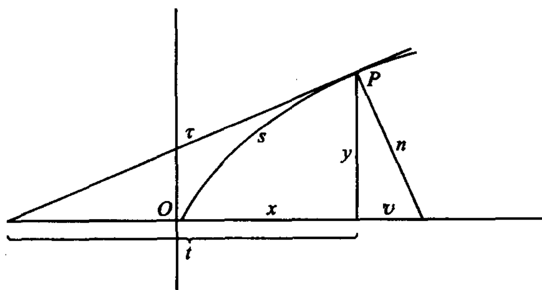


图 12.1 曲线的相关的量:x 为横坐标,y 为纵坐标,s 为弧长,t 为二次切距,τ 为切线,n 为法线,v 为次法距.

本章中我们将首先探讨各种用来作切线和求极值的方法,然后是为确定面积和体积发展的方法,第三是幂级数的思想,它在扩大可以运用面积技巧的曲线的范围方面被证明是极为有用的,第四是确定曲线长度的方法,这曾经被笛卡儿说成是无法完成的事情.随后我们将考虑首先能够将这些各种各样的思想结合成一个统一整体的数学家 —— 主要是牛顿和莱布尼茨 —— 的工作,最后我们会研究在微积分新领域中最早的教科书的内容.

12.1 切线和极值

1615年,开普勒(Johann Kepler,1571—1630)写下了他的《酒桶的新立体几何》,其中他说明了奥地利的葡萄酒商人有一种确定给定的酒桶中还剩下多少酒的相当准确的方法.作为这一对各种立体形状所作研究的一部分,他证明了能够内接于一给定球体的最大的平行六面体是正方体.事实上,他实际对从 1 到 20 的所有整数高度列出了内接于一半径为 10 的球体的平行六面体的体积.他因此非常清楚在大约为 1540 的最大值附近,随高度的微小变化体积变化得很小:"在最大值附近,在两端的减少开始变得难以察觉."[3]

12.1.1 费马的等同法

17 世纪 20 年代晚期,费马得以将开普勒的思想转化为一种算法,但他考虑这一问题的动因却是来自对韦达在多项式的系数与根的关系方面工作的研究:"我在思索韦达的方法时 …… 当时在细究它在发现方程的结构方面的应用时,一种可以用在寻找最大和最小值上的新方法涌上心头,通过这种方法,曾经困扰古代和现代几何的同条件有关的一些疑惑最容易被消除."[4]

韦达曾经证明过方程 $bx - x^2 = c$ 的两个根 x_1, x_2 的和为 b,方法是令 $bx_1 - x_1^2$ 与 $bx_2 - x_2^2$ 相等,并除以 $x_1 - x_2$.方程 $bx - x^2 = c$ 源自"分长度为 b 的线段为两部分,使得它们的积为 c"这样一个几何问题.费马从欧几里得那里得知 c 的最大可能值为 $b^2/4$,并且对于任何小于该最大值的数,x 有和为 b 的两个可能的值.但 c 在接近它的最大值时会发生什么事呢?该几何场景使得费马非常清

楚即便对这一最大值来说,方程仍然有值相等的两个解,$x_1 = b/2, x_2 = b - x_1 = b - b/2 = b/2$. 这一洞见使得费马得到了最大化多项式 $p(x)$ 的方法:令 $p(x_1) = p(x_2)$. 然后除以 $x_1 - x_2$ 来寻找系数和多项式任意两个根的关系. 最后,令这两个根彼此相等并求解.

费马从 $bx_1 - x_1^2 = bx_2 - x_2^2$ 得出 $b = x_1 + x_2$ 这一对任两个根都成立的方程. 令 $x_1 = x_2 (= x)$ 得 $b = 2x$. 因此最大值在 $x = b/2$ 时出现. 同样,为使 $bx^2 - x^3$ 最大化,费马令 $bx_1^2 - x_1^3 = bx_2^2 - x_2^3$ 得到 $b(x_1^2 + x_2^2) = x_1^3 - x_2^3$ 并推得 $bx_1 + bx_2 = x_1^2 + x_1x_2 + x_2^2$. 他接着令 $x_1 = x_2 (= x)$ 并确定出 $2bx = 3x^2$,从而得到结论:$x = 2b/3$ 给出最大值. 费马的过程这里带来两个明显的问题. 他为什么舍弃 $x = 0$ 这个解,以及他如何得知他选取的解给出的是最大值而不是最小值?答案对费马是简单的. 问题是同体积有关的几何问题. 所以 $x = 0$ 的答案没有意义,并且正的答数显然给出最大值. 但更一般地,费马没有考虑当他的方法给出的最后方程有两个或者更多个解的时候会发生什么事?在所有具体的情况下,他都简单地诉诸场景的几何学. 费马的方法还产生另一个问题. 人们如何能在除以 $x_1 - x_2$ 之后再令它的值为0?对费马来说,几何场景说明即便在两个根的差值为0时它们仍然是可以区分的. 因此,他从没有感觉到他在除以 0. 他简单地以为用韦达的方法得出的关系是绝对普遍的(例如,$x_1 + x_2 = b$),因此对变量的任一特殊值成立,即便是最大值处的值.

费马的确意识到,如果多项式 $p(x)$ 比较复杂时,除以 $x_1 - x_2$ 可能会非常困难. 因此他为了避免这一点,对他的方法作了改进. 他把两个根写成 x 与 $x + e$ 而不再将它们考虑为 x_1 和 x_2. 然后,在令 $p(x)$ 与 $p(x + e)$ 相等后 —— 费马实际使用的是拉丁文的**等同**(adequate),这是他从丢番图著作中读来的 —— 他只需要用 e 或者它的幂去除. 在得到的表达式中他剔除掉所有含有 e 的项得到可以找到最大值的方程. 例如,使用他原始的例子 $p(x) = bx - x^2$,费马将 $bx - x^2$ 与 $b(x + e) - (x + e)^2 = bx - x^2 + be - 2ex - e^2$ 相等同.(我们将其记为 $bx - x^2 \approx bx - x^2 + be - 2ex - e^2$.)消掉公共项,得到 $be \approx 2ex + e^2$,再除以 e,得到 $b \approx 2x + e$. 剔除含有 e 的项,费马得到了他知道的结果:$x = b/2$. 在他对这一过程的描述中 —— 该描述或许完成于1630年之前但在1636年末才传到巴黎 —— 费马写道:"我们很难期望一个比这更普遍的方法了."[5]

在同一份文件中,费马说明了他的等同法如何可以被稍加修改用来确定曲线的切线,特别是抛物线的切线. 因为费马在他发明解析几何以前就发明了这一方法,他使用了抛物线的一种几何描述. 但是在 1638 年,一旦使用代数方程而不是几何性质定义曲线的可能性被打开,费马可以更容易地解释他的方法.(事实上,笛卡儿强烈地批评了他的几何解释.)为在点 B 处像用现代符号表示为 $y = f(x)$ 的曲线作切线,在切线上任选一点 A 并向轴作垂线 AI 及 BC(图 12.2). 费马的想法是将 FI/BC 与 EI/CE 相等同,此处 F 是 AI 与曲线的交点. 如果 $CI = e, CD = x, CE =$ 次切距 t,这一等同可以写为

$$\frac{f(x+e)}{f(x)} \approx \frac{t+e}{t},$$

图 12.2　确定次切距的费马方法.

或者 $tf(x+e) \approx (t+e)f(x)$. 通过运用他消除公共项、以 e 除然后再剔除包含 e 的剩余项的方法，费马能够计算出确定切线的 t 和 x 的关系. 例如，如果曲线是抛物线 $f(x)=\sqrt{x}$，则费马的方法给出 $t\sqrt{x+e} \approx (t+e)\sqrt{x}$ 这一等同式. 两边取平方并化简得：$t^2e \approx 2etx + e^2x$. 如果我们用 e 除再剔除仍然含有 e 的项，我们得到结果 $t=2x$. 这当然是阿波罗尼乌斯的结果(命题 Ⅰ－33)：在一点处抛物线的次切距是横坐标的两倍.

在回应笛卡儿提出的挑战中，费马为处理用形式 $f(x,y)=0$ 表达的曲线改进了他的方法，而且事实上通过向笛卡儿说明如何使用他的方法找出曲线 $x^3+y^3=pxy$ 的切线证明了他的方法的有效性，这一曲线是笛卡儿向他提出的.[6]

12.1.2　笛卡儿和法线法

笛卡儿批评费马的一个原因是费马独立于他这位大哲学家发现了同样的新数学. 而且笛卡儿对他自己在任一点向曲线作法线的发现感到十分自豪，自然人们也能用这一方法轻易地确定切线. 正如笛卡儿在他的《几何学》中所写："我敢说，这不仅是我所了解的几何学中最有用最一般的问题，甚至也是我所期望了解的几何学中最有用最一般的问题."[7]

笛卡儿意识到圆的半径总是圆周的法线，从中他得出了画法线的主意. 在给定点同曲线相切的圆的半径必然也是该曲线的法线. 作与曲线相切的圆需要同费马相似的主意，即如果圆的确与曲线相切，在给定点附近圆与曲线的两个交点将变成一个. 为在由曲线 $y=f(x)$ 上一点 C 执行这一程序，假定 P 是所求圆的中心，在通过点 P 的轴上任取一点 A，令 $CP=n$ 且 $PA=v$(图 12.3). 如果 $C=(x,y)$，则 $PM=v-x$ 并且圆的方程为 $n^2=y^2+(v-x)^2$. 笛卡儿然后用这一方程确定 v，v 再确定点 P. 如他所写："如果点 P 满足要求的条件，以 P 为圆心通过点 C 的圆将与曲线 CE 相切而不是相割；但如果这一点 P 比它应该在的位置哪怕稍微近或者远一点，该圆则与曲线相割在 C 以及另外一点. 现在如果该圆同曲线也相割在点 E，方程……必定有两个相等的根……但点 C 和点 E 取得越近，两个根相差得也越小；且当两个点重合时，根精确地相等."[8] 换言之，因为点 P 是相切圆的圆心，确定圆与曲线 $y=f(x)$ 交点的方程 $[f(x)]^2+v^2-2vx+x^2-n^2=0$ 必定有重根. 如同笛卡儿从他对方程的根的研究中所知道的，这意味着多项式有因子 $(x-x_0)^2$，这里 x_0 是重根. 此时令 $[f(x)]^2+v^2-2vx+x^2-n^2=(x-x_0)^2q(x)$，并使 x 的同次项的系数相等，笛卡儿可以用 x_0 表达出 v.

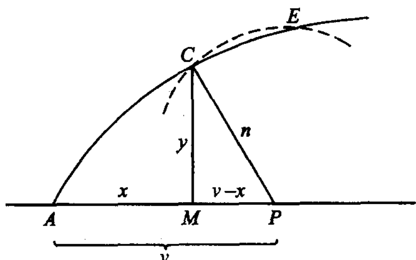


图 12.3　笛卡儿求法线的方法.

像在《几何学》中常有的，笛卡儿为他的程序提供了非常复杂的例子. 因此我们提出一个简单的例子来说明他的方法，即，确定抛物线 $y=x^2$ 在点 (x_0,x_0^2) 处的法线. 在这一情况下，有重根的多项式为 $(x^2)^2+v^2-2vx+x^2-n^2$. 因为这是一个四次多项式，它一定等于 $(x-x_0)^2q(x)$，此时 $q(x)$ 为二次. 因此 $x^4+x^2-2vx+v^2-n^2=(x-x_0)^2(x^2+ax+b)$ 或者 $x^4+x^2-2vx+v^2-n^2=x^4+(a-2x_0)x^3+(b-2x_0a+x_0^2)x^2+(ax_0^2-2bx_0)x+bx_0^2$.

等同系数，得

$$a-2x_0=0,$$
$$b-2x_0a+x_0^2=1,$$

$$a x_0^2 - 2bx_0 = -2v,$$
$$b x_0^2 = v^2 - n^2.$$

通过取 $a = 2x_0, b = 2x_0a - x_0^2 + 1$ 在前三个方程中解 v,得 $v = 2x_0^3 + x_0$ 是所求点 P 的(水平)坐标.(因为 v 确定 n,第四个方程是不必要的.)因为笛卡儿仅仅对作法线有兴趣,在确定了点 P 之后他就中断了程序.但我们进一步注意到法线的斜率是:

$$\frac{-y_0}{v - x_0} = \frac{-x_0^2}{2x_0^3} = \frac{-1}{2x_0},$$

因此切线的斜率是 $2x_0$,一个熟悉的结果.

12.1.3 胡德和斯卢兹的算法

在 17 世纪 30 年代晚期,吉尔斯·波索·德·罗伯华(Gilles Persone de Roberval,1602—1675)发现了一种将曲线看作是由动点产生的确定切线的动力学方法.但他的方法依赖于对曲线的几何描述,因而不能满足人们对确定切线的简单代数算法的需要.费马的程序尤其是笛卡儿的程序导出了如此复杂的代数以至于这些方法也不能提供所要求的计算的便捷.但对这些方法的研究使得另外两名数学家约翰·胡德(Johann Hudde,1628—1704)和勒内·弗朗索瓦·德·斯卢兹(René Francois de Sluse,1622—1685)在 17 世纪 50 年代发现了更简单的算法.

胡德是范·舒滕的一个学生,他像德·威特一样积极参加了荷兰的政治生活.他对数学的贡献完成于17世纪50年代晚期,他的两篇论文出现在范·舒滕1659年版的笛卡儿的《几何学》中.在《论最大值和最小值》中,胡德描述了他简化确定多项式方程重根所需计算的算法,这类计算对求法线的笛卡儿方法是必需的.胡德法则断言,如果一个多项式 $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ 有一重根 $x = \alpha$,并且 $p, p + b, p + 2b, \cdots, p + nb$ 是一算术数列,则多项式 $pa_0 + (p + b)a_1x + (p + 2b)a_2x^2 + \cdots + (p + nb)a_nx^n$ 也有根 $x = \alpha$,胡德对他的法则只给出了证明的大概.用现代的术语说,新的多项式可以表示为 $pf(x) + bxf'(x)$.胡德的结果立即得出,因为如果 $f(x)$ 有重根,则 $f'(x)$ 有同样的根.虽然胡德的法则允许任意选择算术数列,他最常用的是取 $p = 0, b = 1$ 时的数列.在这一情况下,新多项式是 $xf'(x)$,这一结果有助于显示出我们今天所说的导数在计算上的重要性.

作为法则的第一个例子,考虑确定抛物线 $y = x^2$ 法线的问题,这需要找到系数 v 和多项式 $x^4 + x^2 - 2vx + v^2 - n^2$ 的重根 x_0 的关系.利用 $p = 0, b = 1$ 时的胡德法则得到新的多项式 $4x^4 + 2x^2 - 2vx = 0$ 或者 $4x^3 + 2x - 2v = 0$.因为 x_0 是这个方程的一个解,可以像从前一样得出 $v = 2x_0^3 + x_0$,所以切线的斜率为 $(v - x_0)/x_0^2 = 2x_0$.对该例的简单推广使得说明 $y = x^n$ 在点 (x_0, x_0^n) 处的切线的斜率为 nx_0^{n-1} 成为可能,这是一个应用笛卡儿方法极难得到的结果.

胡德还用他的法则确定极值,这要用到费马的主意,即如果一个多项式 $f(x)$ 有极值 M,则多项式 $g(x) = f(x) - M$ 有重根.例如要最大化 $x^2(b - x)$,对多项式 $-x^3 + bx^2 - M$ 利用 $p = 0, b = 1$ 时的胡德法则.新的多项式为 $-3x^3 + 2bx^2 = 0$,它的非零根 $x = 2b/3$ 给出所求的最大值.胡德还使用他的法则确定由形如 $f(x, y) = 0$ 的方程确定的曲线的切线,但斯卢兹对该情形给出了更为简单的算法.

斯卢兹生于现在属于比利时的里格并且一生的大部分时间在这里度过,像胡德一样,他在数学上也没有多少时间.但他同遍及欧洲的数学家进行了广泛的通信.他的确定由多项式方程 $f(x, y) = 0$ 给出的曲线的次切距(当然还有切线)的方法可能发现于 17 世纪 50 年代但直到 1673 年才以书

面的形式出现在给英国的亨利·奥尔登堡(1615—1677)的一封信中.算法从消去常数项开始.然后将所有含 x 的项留在左边,而将所有含 y 的项在适当的符号变换后移到右边.因此所有既含 x 又含 y 的项此时将出现在方程的两边.再然后将右边的各项乘以该项 y 的指数而将左边各项乘以该项的 x 的指数.最后,将左边各项中的一个 x 用 t 置换并就 t 求解得到的方程.例如,所给方程是 $x^5 + bx^4 - 2q^2y^3 + x^2y^3 - b^2 = 0$,除掉常数项并将所有含 y 的项移项得 $x^5 + bx^4 + x^2y^3 = 2q^2y^3 - x^2y^3$.用那些适当的指数相乘并将左边每项中的一个 x 换成 t,得 $5x^4t + 4bx^3t + 2txy^3 = 6q^2y^3 - 3x^2y^3$.因此次切距 t 由下式给出:

$$t = \frac{6q^2y^3 - 3x^2y^3}{5x^4 + 4bx^3 + 2xy^3}.$$

而切线的斜率为:

$$\frac{y}{t} = \frac{5x^4 + 4bx^3 + 2xy^3}{6q^2y^2 - 3x^2y^2}.$$

用现代的术语说,很容易看出斯卢兹计算了

$$t = -\frac{xf_y(x,y)}{f_x(x,y)} \quad \text{或者} \quad \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{f_x(x,y)}{f_y(x,y)}.$$

但斯卢兹没有书面说明他的方法的正当性也没有暗示他是如何发现该方法的.最好的猜测是他从对许多例子的研究中推广而来的.无论如何,胡德和斯卢兹法则的重要性在于他们提供了通用的算法,人们利用这些算法可以按部就班地作出多项式方程给出的曲线的切线.从此再没有必要为每一种特定的曲线来发展一种特殊的技巧了.每个人现在都能确定切线了.

12.2 面积和体积

希腊和伊斯兰的数学家都能够确定由曲线或者曲面围成的某些区域的面积或体积.但现有的文字资料一般都只给出结果和一个基于穷竭法的证明.这些结果在如何确定由现在可以研究的许多新曲线围成的面积或者通过绕平面中直线旋转这些曲线产生的立体区域的体积方面不能给17世纪的数学家们提供多少线索.从希腊时期传下来的惟一清晰的观念是给定区域需要以某种方式分割成单个区域的面积或体积已知的细小的区域.

12.2.1 无穷小量和不可分量

回忆一下开普勒在他发现行星运动定律时使用了加总小区域的程序.而且在他的《新立体几何》中,他计算了半径为 AB 的圆的面积,他首先注意到"圆周……有点这么多的部分,即无穷多个部分;每个部分都可视为腰为 AB 的等腰三角形的底,因此圆面积中有无数个三角形,它们的顶点都在圆心 A."[9] 开普勒然后将圆周展为一条直线,他在直线上的每一点,"一个接一个放",都放上同圆内三角形相等的三角形,所有三角形的高都为 AB(图 12.4).由此得出三角形 ABC 的面积"因为由所有这些三角形组成,将等于圆的所有扇区从而等于包含所有扇区的圆的面积."所以,圆的面积等于半径与周长的积

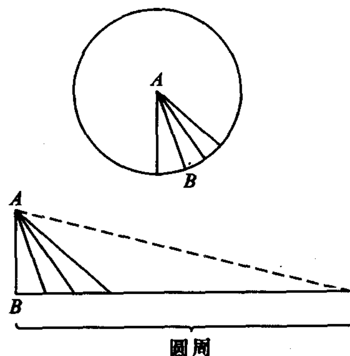


图 12.4 确定圆面积的开普勒方法.

的一半;或者,如开普勒表述的,圆的面积与直径平方之比等于 11 : 14. 同样,开普勒通过把环面切成"无数个很薄的圆盘"[10] 计算了环面的体积,每个圆盘越靠近中心越薄,越靠外越厚. 但开普勒从未声称他的方法是严格的,只提到"如果我们不畏惧阅读阿基米德著作的痛苦,我们从中就可以得到绝对的在各方面都完善的证明."[11]

开普勒对"非常薄"的圆盘和非常小的三角形的应用展示了后来人们所称的**无穷小方法**. 相反,伽利略则使用了**不可分量方法**,在这种方法中一个给定的几何体被视为由比它低一维的对象所构成. 例如,像阿基米德一样,他视平面图象为由线构成而立体图形由面构成. 他也不相信他需要阿基米德的论证来说明它们应用的正当性. 如他所写:

"我说直线由点和不可分量的连续体构成是最真实而且是必要的 ……,认清楚这个连续体可以分成部分,永远可分,原因仅仅是它是由不可分量组成的. 因为,如果分割和下一层分割可以永远进行下去,众多的部分必定是不能穷尽的,因此部分在数量上是无限的,否则下一层分割将会到头,而如果它们是无限的,它们必定是没有大小的,因为无穷多有大小的部分将构成无穷大."[12]

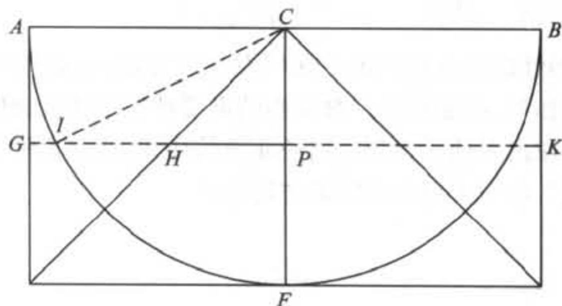


图 12.5 伽利略的从圆柱中切出的汤碗.

但无穷有一些奇怪的性质,伽利略在他对"汤碗"体积的计算中说明一个点和一条直线可能存在的相等. 我们从一个放置在一个圆柱体中的半球 AFB 开始并同时考虑一个顶点在直径 AB 上的点 C 而底与 AB 相等的圆锥(图 12.5). 如果我们从实心圆柱中除去球,剩下的就是称为汤碗的实心区域. 为计算汤碗的体积,伽利略考虑沿 GK 的水平截面(对图中所有立体的竖直截面). 我们有 $IC^2 = IP^2 + PC^2$. 但 $IC = AC = GP$;所以 $GP^2 = IP^2 + PC^2$. 但 $PC = PH$,且所以 $GP^2 = IP^2 + PH^2$,或者 $GP^2 - IP^2 = PH^2$. 因为圆锥和汤碗是绕中心轴 CF 旋转各线段生成的,且圆和它们直径上的正方形相似,由此得出,汤碗的截面在面积上等于圆锥的截面. 根据早先被希罗(Heron)用过的原理,在对应的高度上具有相等截面积的两个立体具有相同的体积,伽利略得出汤碗的体积等于圆锥的体积的结论. 但伽利略更为关心的不是体积,而是以上证明的在每一个平面上成立的等式在图形的顶端也必定成立的事实,在这一情况下圆锥等于一个点而汤碗等于一个完整的圆. 因此,一个点同一条线相等. 如同伽利略所说,"假如它们是等量留下来的最后残余,那么为什么这些不能被称为相等呢?"[13]

人物小传	波那番图拉 · 卡瓦列里(1598—1647)(Bonaventura Cavalieri)
	卡瓦列里是在比萨担任一个小的修道会的成员时开始研究数学的,也是在这里他开始了同伽利略长期的通信,通信几乎一直持续到后者去世. 也许是通过后者的影响,他在 1629 年在波罗尼亚大学获得教授席位并在每三年一次的改选中成功地获得连任直到他自己去世. 除了在正文中提到的那些著作,卡瓦列里还出版了许多有关数学的书籍,包括一本星象学的著作,他还研究透镜和镜. 但他的声誉是建立在《几何学》中的不可分量方法上的,该著作广为人知,但因为它的难度,也许很少有人研究.

第一个发展出不可分量完整理论的人是波那番图拉·卡瓦列里(Bonaventura Cavalieri, 1598—1647),伽利略的一个追随者.他在他1635年的《借助连续量的不可分量用新方式提出的几何学》和他1647年的《六个几何练习题》中对该理论作了详细阐述.卡瓦列里著作的核心概念是一个平面图形 F 的"所有的线",记为 $\mathcal{O}_F(l)$.卡瓦列里用它表示当一个垂直平面与自身平行着从给定图形的一端移向另一端时与该平面图形的交线的集合.交线是直线,而卡瓦列里的整个工作处理的正是这个被视为一个量的这些直线的集合.卡瓦列里的直线在某种意义上构成了给定的图形,但他仔细地把 $\mathcal{O}_F(l)$ 同 F 本身区分开来.他也能够通过考虑给定图形的更高维的对象来推广他的思想,如"所有的正方形"或者所有的"立方体".我们可以想象一个三角形的"所有的正方形",例如像表示棱锥那样,它的每一个横截面都是边长为三角形中一特定直线长度的正方形.

卡瓦列里计算的基础是今天以卡瓦列里原理闻名的结果,该原理我们在希罗和阿基米德的著作中已经看到过:"如果两个平面图形的高相等并且由平行于底并与底距离相等的直线所截得的线总是成相同的比例,则平面图形的面积也成这一比例."[14] 卡瓦列里用叠合的论证法证明了该结果.由此可得如果两图形 F 和 G 的对应直线间成固定的比例,则 $\mathcal{O}_F(l):\mathcal{O}_G(l) = F:G$.例如,假设长为 a 宽为 b 的矩形 F 被它的对角线分为两个三角形 T, S(图12.6).因为三角形 T 中每一线段 BM 对应一条而且仅一条三角形 S 中的直线 HE,则 $\mathcal{O}_T(l) = \mathcal{O}_S(l)$.另一方面,因为矩形的每一线段 BA 都由一条三角形 S 中的线段和一条三角形 T 中的线段构成,则 $\mathcal{O}_F(l) = \mathcal{O}_T(l) + \mathcal{O}_S(l)$.由此得出 $\mathcal{O}_F(l) = 2\mathcal{O}_T(l)$,或者矩形的所有的直线是三角形的所有的直线的一倍.用现代的概念,这一结果等价于 $ab = 2\int_0^b \frac{a}{b}t\,\mathrm{d}t$ 或者更简单地,等价于

$$b^2 = 2\int_0^b t\,\mathrm{d}t.$$

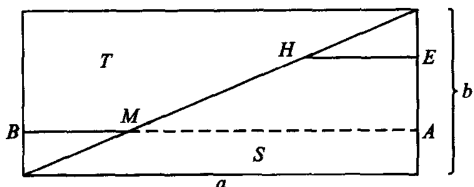


图12.6　卡瓦列里的一个三角形和矩形的"所有直线"的方法.

卡瓦列里同样可以证明矩形 F 的"所有的正方形"是每个三角形的"所有的正方形"的三倍,或者用现代的概念,为

$$a^2b = 3\int_0^b \frac{a^2}{b^2}t^2\mathrm{d}t \quad \text{或} \quad b^3 = 3\int_0^b t^2\mathrm{d}t.$$

1647年前,他对一些高次幂证明了相似的结果并且能够推出内接在一个矩形里的高次抛物线 $y = x^k$ 下的面积是该矩形的 $\frac{1}{k+1}$ 倍,或者

$$\int_0^b x^k\mathrm{d}x = \frac{1}{k+1}b^{k+1}.$$

同一时期费马、帕斯卡、罗伯华以及托里拆利也发现了这个结果.

12.2.2　托里拆利和无限长的立体

伊凡哥里斯塔·托里拆利(Evangelista Torricelli, 1608—1647),伽利略的另一个追随者,也用不可分量工作,但他告诫说不加批判地使用不可分量可能得出悖论.例如,假定一个人使用相互垂直

的不可分量计算矩形 $ABCD$(图 12.7) 中的两个三角形的面积. 在这种情况下,因为直线 FE 与直线 EG 的比始终等于直线 AB 与 BC 的比,似乎可以得出三角形 ABD 与三角形 DBC 的面积也成相同的比例,一个荒谬的结果. 托里拆利解决该悖论的方法实质上是转到无穷小量上,即,认为"不可分"线段实际上有一个厚度. 在这一具体例子中,竖直线段比水平线段要厚,比例是 $AB : BC$,所以如果人们把所有的线段放在一起,三角形 ABD 与三角形 DBC 事实上有相同的面积. 虽然托里拆利的许多工作在他生前并未发表,但它确实在他自己学生的著作中在意大利传播. 因此,我们知道他解决了确定曲线 $y = x^{m/n}$ 下面积以及确定该曲线的切线这样的问题. 有趣的是,同他的许多同时代人不同,他通常用双归谬法对他的结果给出了完整的经典证明.

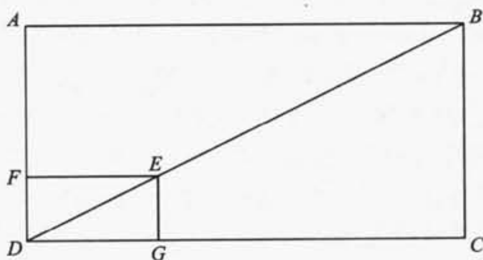


图 12.7　托里拆利利用不可分量的理论.



图 12.8　意大利邮票上的托里拆利.

> **人物小传**
>
> **伊凡哥里斯塔 · 托里拆利(1608—1647)(Evangdista Torricelli)**
>
> 托里拆利(图 12.8) 在罗马跟随伽利略的一个学生波耐戴托 · 卡斯特里(1578—1643) 学习数学,到了 1641 年得以在伽利略在阿切里的家里跟随伽利略本人学习. 他在那里一直呆到伽利略去世并在不久后被任命为原来由伽利略担任的托斯卡纳区大公的数学家和哲学家. 托里拆利在佛罗伦萨度过余生,继续伽利略在运动方面的工作并为威力更大的望远镜磨制透镜. 他最出名的也许因为在 1643 年发现了气压计原理. 1647 年,他死于伤寒症.

然而托里拆利最令人惊异的发现是在1643年宣布的. 他说明了绕 y - 轴旋转双曲线 $xy = k^2$ 形成的无限长立体从 $y = a$ 到 $y = \infty$ 段的体积是有限的,并且实际上它与半径为 k^2/a、高度为 a 的圆柱的体积之和等于半径为此双曲线的半 - 直径 $AS = \sqrt{2}k$、高为 k^2/a 的圆柱的体积(图 12.9). 托里拆利使用了与今天讲授的圆柱壳方法类似的方法,但用了与他的朋友卡瓦列里的线相类似的不可分量的方法来表述. 首先,他说明了任何内接在他的无限双曲立体中的圆柱的侧表面积都等于半径为 AS 的圆的面积,例如圆柱 $POMN$. (用现代的术语说,这意味着 $2\pi x(k^2/x) = \pi(\sqrt{2}k)^2$.) 然后,他注意到该无限立体(包括它底端的圆柱) 可以被看作由所有那些每个对应着组成圆柱 $ACHI$ 的一个圆的圆柱表面构成. 由此可得无限圆柱的体积等于圆柱 $ACHI$.

托里拆利写道:"尽管这一立体有无限的长度,但我们考虑的所有圆柱面都没有无限的长度而是有有限的长度,这似乎是难以置信的."[15] 因为他相信他的结果是"难以置信的",但他决定给出另一个证明来加强他的结果,这个证明用的是穷竭法.

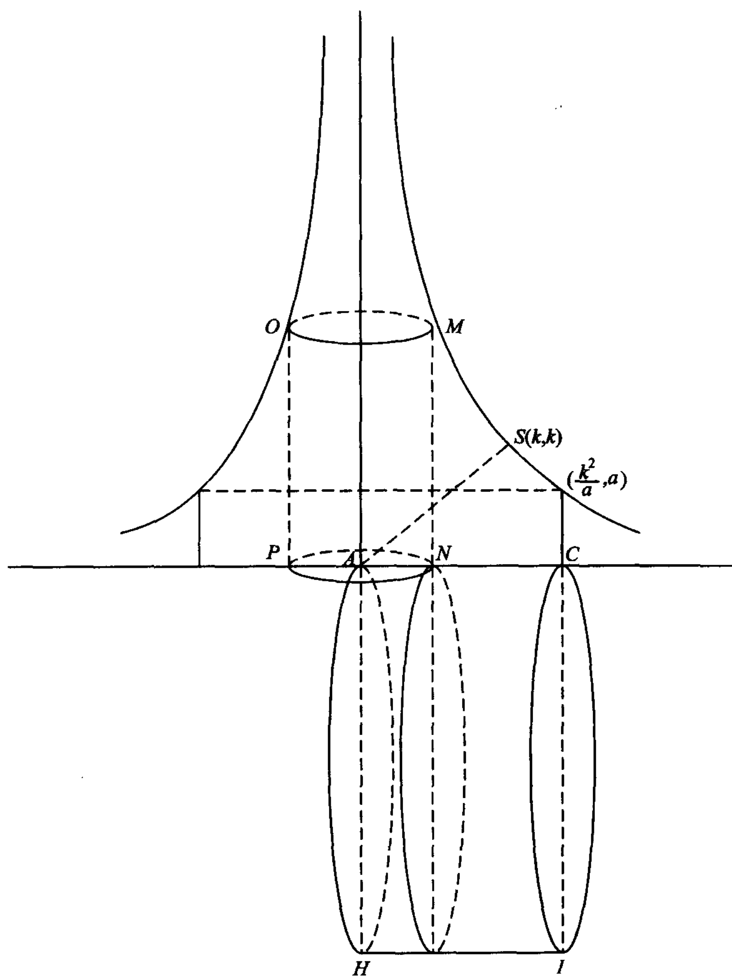


图 12.9 托里拆利的无限双曲体.

12.2.3 费马和抛物线与双曲线下的面积

在 1636 年 9 月 22 日给罗伯华的一封信中,费马声称他能够弄方"无穷多由曲线组成的图形",特别是他能够计算任何高次抛物线 $y = px^k$ 下的区域的面积. 他进一步写道:"我不得不遵循一条与阿基米德求抛物线面积所用方法不同的路线,如果用后者的方法我将永远解决不了问题."[16] 费马这里回忆起阿基米德在求面积时利用了三角形,费马自己将使用更简单的图形. 罗伯华 10 月作了回复,他声称他利用自然数的幂和公式也发现了同样的结果:"平方数的和永远大于以最大的平方数的底为底的三次方的三分之一,同样的平方数的和在除去最大的平方数后小于以最大的平方数的底为底的三次方的三分之一;立方数的和永远大于以最大的立方数的底为底的四次方的四分之一,同样的立方数的和在除去最大的立方数后小于以最大的立方数的底为底的四次方的四分之一,以此类推."[17] 换言之,求由抛物线 $y = px^k$、x 轴和一条给定的竖直直线围成的区域的面积取决于以下公式

$$\sum_{i=1}^{N-1} i^k < \frac{N^{k+1}}{k+1} < \sum_{i=1}^{N} i^k.$$

考虑一下 $y = px^k$ 在区间$[0, x_0]$上的图象就很容易看出为什么这一公式是关键的.将底部区间分成 N 个相等的子区间,每一个的长度为 x_0/N,在每一子区间上作一高为右端点的 y - 坐标的矩形(图 12.10).这 N 个外接矩形的面积之和则为

$$p\frac{x_0^k}{N^k}\frac{x_0}{N} + p\frac{(2x_0)^k}{N^k}\frac{x_0}{N} + \cdots + p\frac{(Nx_0)^k}{N^k}\frac{x_0}{N} = \frac{px_0^{k+1}}{N^{k+1}}(1^k + 2^k + \cdots + N^k).$$

同样,人们可以计算高为相应区间左端点的 y - 坐标的内接矩形的面积和.如果 A 是曲线下 0 与 x_0 之间部分的面积,则

$$\frac{px_0^{k+1}}{N^{k+1}}(1^k + 2^k + \cdots + (N-1)^k) < A < \frac{px_0^{k+1}}{N^{k+1}}(1^k + 2^k + \cdots + N^k).$$

这一不等式的外项之差不过是最右边一个外接矩形的面积.因为 x_0 与 $y_0 = px_0{}^k$ 是定值,只要取 N 为足够大就可以使该差小于任何指定的值.从罗伯华引用的不等式可以推出面积 A 和值$\frac{px_0^{k+1}}{k+1} = \frac{x_0y_0}{k+1}$都在两个差趋于 0 的量间"夹挤".因此费马(以及罗伯华)发现 $A = \frac{x_0y_0}{k+1}$.

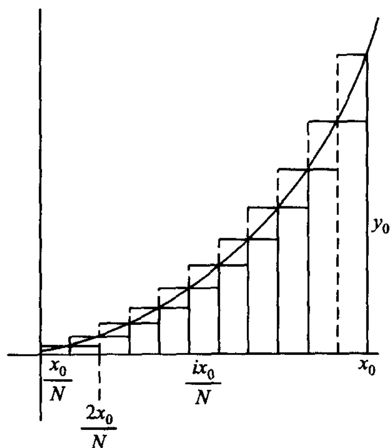


图 12.10　按费马和罗伯华的方法求 $y = px^k$ 下方的面积.

那么,明显的问题是这两个人是如何发现幂和公式的,该公式实际上在600年前已经为伊本·阿尔·海塞姆所知.费马声称他有一个"精确的证明"并对罗伯华有没有表示怀疑.实际上,像在费马工作中典型的那样,我们所能看到的只有他关于数、棱锥数和其它作为列出现在帕斯卡三角形中的数的笼统命题:"最后一边的数乘以上一边等于三角形数的两倍.最后一边乘以上一边中的三角形数等于棱锥数的3倍.最后一边乘以上一边的棱锥数等于重三角形数的 4 倍.以此类推以至无穷".[18] 我们记为

$$N\binom{N+k}{k} = (k+1)\binom{N+k}{k+1}$$

的费马的命题等价于帕斯卡的第 12 个结果.利用帕斯卡三角形的性质,对于按顺序的每一个 k(从 $k = 1$ 开始)不难推出 k 次幂的和的明确的公式.该公式的形式是

$$\sum_{i=1}^{N} i^k = \frac{N^{k+1}}{k+1} + \frac{N^k}{2} + p(N),$$

这里,$p(N)$ 是次数不高于 k 的关于N 的多项式.对 $p(N)$ 形式的仔细研究则可以使我们推出罗伯华的不等式.

我们不知道费马果真证明了他说的普遍结果还是仅仅在对若干 k 的值作了验算后就假定它对所有的值都成立;我们也不知道费马是如何推导出整数的幂和公式的.也许,费马不知道乌尔姆的计算大师约翰·佛尔哈伯(Johann Faulhaber,1580—1635)的工作,他在 1631 年推导出了直到 17 次幂的明确的整数幂和公式.[19] 并且帕斯卡本人,在 1654 年写作时,可能也不知道费马的结果,他从他的三角形的性质中给出了幂和公式的明确推导,他还写道:"那些对不可分量的学说稍微有所了解的人将会看到,人们可以利用这一结果确定曲线图形的面积.该结果立即使得人们能够求所有类型的

抛物线和无数种其它曲线围成的面积."[20]

无论如何,费马对他求面积的方法并不完全满意,因为它仅适用于高次抛物线.他看不出如何使它适用于形如 $y^m = px^k$ 的曲线或者形如 $y^m x^k = p$ 的"高次抛物线".用现代的术语说,这一求 $y = px^k$ 下面积的方法仅适用于 k 为正整数的情形.费马想得到一个适用于 k 为任意正或者负有理数情形的求面积的方法.尽管他在1658年前后的《求积论》中才宣布了这种方法,似乎很明显,他在17世纪40年代就发现了这一新程序.

在确定 $x = x_0$ 之右 $y = px^{-k}$ 下面积的问题中应用他早先的方法需要把或者 x - 轴或者从0到 $y_0 = px_0^{-k}$ 间的直线段 $x = x_0$ 分为有限多个区间并把内接和外接矩形的面积求和.但应用后一过程将使费马得到作为内接和外接矩形之差的无限矩形,该矩形能否被弄成任意小压根是不清楚的.另一方面,没有办法将无穷的 x - 轴分成有限多个最终可以分到任意小的区间.费马解决这一难题的办法是将 x - 轴分成无限多个长度不相等而是成几何级数的区间,再用对这种数列求和的已知公式加上这无穷多个矩形的面积.

费马从将 x_0 以右的无穷区间在点 $a_0 = x_0, a_1 = \frac{m}{n}x_0, a_2 = \left(\frac{m}{n}\right)^2 x_0, \cdots, a_i = \left(\frac{m}{n}\right)^i x_0, \cdots$ 处进行划分,此处 m 和 $n(m > n)$ 是正整数(图12.11).区间 $[a_{i-1}, a_i]$ 将最终通过将 m/n 取得足够接近于1而被弄到任意小.费马然后在每个小区间上作外接矩形.第一个外接矩形的面积是

$$R_1 = \left(\frac{m}{n}x_0 - x_0\right)y_0 = \left(\frac{m}{n} - 1\right)x_0 \frac{p}{x_0^k} = \left(\frac{m}{n} - 1\right)\frac{p}{x_0^{k-1}}.$$

第二个矩形的面积是

$$R_2 = \left[\left(\frac{m}{n}\right)^2 x_0 - \left(\frac{m}{n}\right)x_0\right]\frac{p}{\left(\frac{m}{n}x_0\right)^k} = \left(\frac{m}{n}\right)\left(\frac{m}{n} - 1\right)x_0\left(\frac{n}{m}\right)^k \frac{p}{x_0^k} = \left(\frac{n}{m}\right)^{k-1} R_1.$$

同样,第三个矩形的面积是 $R_3 = \left(\frac{n}{m}\right)^{2(k-1)} R_1$.

由此推出所有外接矩形的面积的和为

$$\begin{aligned} R &= R_1 + \left(\frac{n}{m}\right)^{k-1} R_1 + \left(\frac{n}{m}\right)^{2(k-1)} R_1 + \cdots \\ &= R_1\left[1 + \left(\frac{n}{m}\right)^{k-1} + \left(\frac{n}{m}\right)^{2(k-1)} + \cdots\right]. \end{aligned}$$

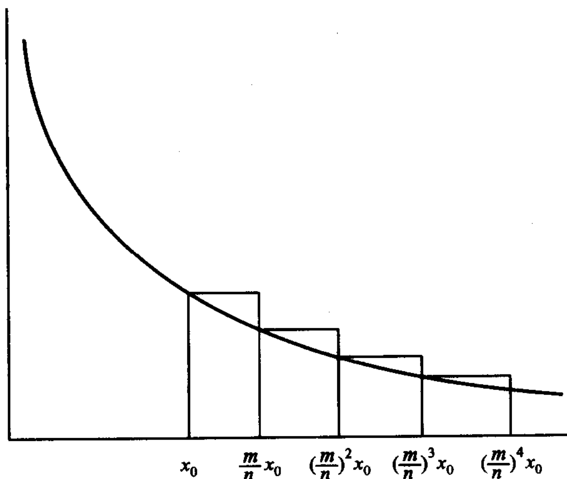


图 12.11　费马确定 $y = px^{-k}$ 下面积的程序.

或者,应用几何级数的求和公式,

$$R = \frac{1}{1-\left(\frac{n}{m}\right)^{k-1}}R_1 = \frac{1}{1-\left(\frac{n}{m}\right)^{k-1}}\left(\frac{m}{n}-1\right)\frac{p}{x_0^{k-1}}$$

$$= \frac{1}{\frac{n}{m}+\left(\frac{n}{m}\right)^2+\cdots+\left(\frac{n}{m}\right)^{k-1}}\frac{p}{x_0^{k-1}}.$$

费马本来可以对内接矩形作同样的计算,但确认这是不必要的. 他使第一个矩形的面积"趋于零";或者,用现代的术语说,通过使 n/m 趋于 1,找到了他的和的极限值. R 的值于是趋于 $\frac{1}{k-1}\frac{p}{x_0^{k-1}}$,于是所求的面积 A 由下式给出

$$A = \frac{1}{k-1}x_0y_0.$$

费马很快注意到这种把轴分为无限个区间的划分也可以被用来求从 $x=0$ 到 $x=x_0$ 处抛物线 $y=px^k$ 下的已知面积. 他简单地把这一有限区间 $[0,x_0]$ 分成无穷多个子区间,从右开始为:

$$a_0 = x_0, a_1 = \frac{n}{m}x_0, a_2 = \left(\frac{n}{m}\right)^2x_0, \cdots, a_i = \left(\frac{n}{m}\right)^ix_0, \cdots,$$

此处 $n<m$ 并进而像上面一样去说明这一面积等于 $A=\frac{1}{k+1}x_0y_0$. 在费马希望解决的另一些情形,即,曲线 $x^ky^m=p$ 和 $y^m=px^k$ 下的面积,该方法必需被稍加改进以避免几何级数中出现分数次幂(补遗 12.1). 但费马的确成功地证明了 $x=x_0$ 以右"双曲线" $x^ky^m=p$ 下的面积为 $\frac{m}{k-m}x_0y_0$,而从 0 到 x_0"抛物线" $y^m=px^k$ 下的面积为 $\frac{m}{k+m}x_0y_0$.

补遗 12.1

费马发明了微积分吗?

在 17 世纪 40 年代中期,费马就已经确定出任何形如 $y=x^k$ 的曲线下的面积(当然,除了 $y=x^{-1}$,费马意识到他的方法对该曲线不适用)并且能够作出这些曲线的切线. 因为他解决了微积分的两个主要问题,至少是在这些重要的特例中,为什么他不应该被当作微积分的发明人?答案必定是费马没有意识到两个问题的逆关系,在某种程度上是因为他不明白我们称为微分和积分的微积分的两种基本运算每种都确定一个新函数,而对于这个新函数,人们可以再度施加这些运算. 今天的学生在看到 $y=x^k$ 的导数是函数 $y'=kx^{k-1}$ 以及从 0 到 x 的 $y=x^k$ 下的面积是函数 $\frac{x^{k+1}}{k+1}$ 后可能会立即意识到逆的性质. 费马意识不到,因为他没有提出会使他作出这一结论的问题. 对费马来说,作切线严格地意味着:求次切距的长度并从曲线上的点向轴上适当的点引直线. 因此,他没有一般地考虑切线的斜率,即我们的导数. 在处理 $y=x^k$ 时,他会得出次切距 t 等于 x/k 而不是切线的斜率等于 kx^{k-1}. 类似地,求曲线下的面积对费马意味着求一个面积等于给定的曲边区域面积的适当的矩形. 换言之,$y=x^k$ 下从 0 到 x_0 的面积等于宽为 x_0 高为 $\frac{1}{k+1}y_0$ 的矩形的面积. 他从未把从一个固定的坐标到一个变量坐标间的面积考虑成确定一个可以表示为新曲线的函数. 因此,尽管费马能够在许多例子中解决微积分的两个基本问题,他没有提出"恰当"的问题. 是别人能够看到费马没能觉察的问题.

12.2.4 沃利斯和分数指数

导出了同费马一样的“积分”公式的另一个数学家是约翰·沃利斯(John Wallis,1616—1703).沃利斯是实际上解释了并前后一致地运用了分数指数的第一位数学家,他从书中了解了卡瓦列里的工作但却无法找到一本卡瓦列里的书.因此,虽然他运用了不可分量,但他在他 1655 年的《无穷算术》中却采取了同卡瓦列里有所不同的研究.为了确定 $x=0$ 与 $x=x_0$ 间 $y=x^2$ 下的面积与外接矩形面积 x_0y_0 的比,他注意到坐标 x 上的对应线段的比为 $x^2:x_0^2$.但因为有无穷多这样的坐标,沃利斯需要计算无穷多前项的和与无穷多后项的和之比.沃利斯将他的坐标取成一个算术数列 0,1,2,…,他希望确定用现代符号表示的下式

$$\lim_{n\to\infty}\frac{0^2+1^2+2^2+\cdots+n^2}{n^2+n^2+n^2+\cdots+n^2}.$$

人物小传	约翰·沃利斯(1616—1703)(John Wallis)
	虽然沃利斯在剑桥大学学的是数学,但他早年的大部分时间是花费在准备作一名牧师上面.无论如何,他在各种科学问题上的兴趣使他积极参加了在 1662 年组建了皇家学会的那一群人在 17 世纪 40 年代举行的最早的非正式聚会.这些每周举行一次的会议致力于“哲学研究”的探讨,内容包括解剖学、几何学、天文学和力学,当时这些学科在英国和大陆都被详尽地探究.沃利斯早期对数学的兴趣在 1647 年前后得以恢复,两年后他被任命为剑桥的萨弗尔教授,这一空缺是因为他的前任在英国内战中选错了立场产生的.在剑桥,沃利斯写下了一系列数学著作,除了《无穷算术》,还包括几何、圆锥曲线和力学方面的论文.

为计算该比,他尝试了不同的情况

$$\frac{0+1}{1+1}=\frac{1}{2}=\frac{1}{3}+\frac{1}{6},$$

$$\frac{0+1+4}{4+4+4}=\frac{5}{12}=\frac{1}{3}+\frac{1}{12},$$

$$\frac{0+1+4+9}{9+9+9+9}=\frac{14}{36}=\frac{1}{3}+\frac{1}{18}.$$

并且,一般地

$$\frac{0^2+1^2+2^2+\cdots+n^2}{n^2+n^2+n^2+\cdots+n^2}=\frac{1}{3}+\frac{1}{6n}.$$

沃利斯的结论是,如果项数是无限多的,就是说,如果直线“充满”了所求的面积,这个比值将准确地为 1/3.在对立方的情形计算出相似的比为 1/4 后,沃利斯作了他称其为“归纳”的跳跃,他达到的结论是,对于任意正整数 k,

$$\frac{0^k+1^k+2^k+\cdots+n^k}{n^k+n^k+n^k+\cdots+n^k}=\frac{1}{k+1},$$

如果有无穷多项的话.

沃利斯的下一步是通过类比将这一结果推广到其它次幂.例如他注意到任意给定的成算术级数的幂,例如 2,4,6,…,对应的面积比的后项也成算术级数,即,3,5,7,… 由此得出如果比的后项为 1,幂的数列的指数必定为 0,就是说,m^0 对任意的 m 都必须是 1.并且,他注意到后项为 3 的二次幂的数列,是由后项为 5 的 4 次幂的数列的平方根所组成,而 3 是 0 次和 4 次幂级数的后项 1 和 5 的

算术平均值.沃利斯接着作了另一个大胆的推广,取项为$\sqrt{0},\sqrt{1},\sqrt{2},\cdots$的级数,这些项是级数 0,1,2,⋯ 的平方根.他确定相应的比的后项应该是 1 和 2 的算术平均值.换言之,比

$$\frac{\sqrt{0}+\sqrt{1}+\sqrt{2}+\cdots\sqrt{n}}{\sqrt{n}+\sqrt{n}+\sqrt{n}+\cdots\sqrt{n}}$$

最终必定等于$\frac{1}{1\frac{1}{2}}=\frac{2}{3}$.另外,这一数列的幂应该是 0 和 1 的算术平均值,即 1/2,或者如沃利斯所说,$\sqrt{x}$ 的指数是 1/2.沃利斯同样总结出$\sqrt[3]{x}$ 的指数一定是 1/3 而$\sqrt[3]{x^2}$ 的指数一定是 2/3,同时它们相应比的后项必定是介于 1 和 2 之间的两个算术平均值,即$1\frac{1}{3}$和$1\frac{2}{3}$.然后,通过把任意的正分数 p/q 的分数次幂定义成p 次幂q 次方根的指数,沃利斯将所有这些推广总结为一个定理:"如果我们取一个量的无穷数列,从一个点或者 0 开始,连续地以任何整数或有理分数次幂的比例增加,则总和与项数同样多的每项等于最大数的级数之比等于这个幂的指数加 1 取倒数."[21]

尽管沃利斯应用这一结果来解决形为 $y=x^{p/q}$ 的曲线的面积问题,就是说,他发现了

$$\int_0^1 x^{p/q}\mathrm{d}x=\frac{1}{\frac{p}{q}+1},$$

但除了指数为 $1/q$ 的情形外他没有证明他的答案是正确的.但是他是一个类比力量的坚定的信仰者.因此,他把他指数的思想推广到负数和无理数并证明了这些指数服从我们熟悉的指数规律.当他试图把他的定理和对形如 $y=x^{-k}$ 的曲线的面积问题的解法推广时,他的方法暴露出了缺陷.他的基本法则告诉他在指数为 -1 的情况对应的比应该是 $1/(-1+1)=1/0$ 而在指数为 -2 的情况比应该是 $1/-1$.假定双曲线 $y=1/x$ 下的面积在某种意义上是 1/0 或者无穷大是合理的,但曲线 $y=1/x^2$ 下的面积为 $1/-1$ 是什么含义呢?因为对于指数 3,2,1,0,对应的比是 1/4,1/3,1/2 和 1/1,这些值形成一个递增的数列,他假定指数 -2 的比 $1/-1$ 应该比指数 -1 的比 1/0 大.但 $1/-1$ 比无穷大大是什么含义呢,沃利斯一直搞不明白.

沃利斯忽略掉这一问题但还是意识到他的方法可以用来求由形如 $ax^{p/q}$ 的项的和给出的曲线下的面积,他进而试图将他的方法推广到用算术方法确定单位圆面积这一更为复杂的问题,问题就是要求曲线 $y=\sqrt{1-x^2}=(1-x^2)^{1/2}$ 下的面积.为了使用他类比论证的技巧,他实际攻击了一个更为普遍的问题,找单位矩形的面积与曲线 $y=(1-x^{1/p})^n$ 在第一象限围成的面积的比.$p=1/2$,$n=1/2$ 时就得到圆的情形,这时比值为 $4/\pi$.沃利斯用他已经知道的方法计算 p 和 n 都是整数情形的比值是相当容易的.例如,如果 $p=2$ 而 $n=3$,从 0 到 1 的 $y=(1-x^{1/2})^3$ 的面积是 $y=1-3x^{1/2}+3x-x^{3/2}$,即 $1-2+3/2-2/5=1/10$.因为单位正方形的面积是 1,这里的比为 $1:1/10=10$.这样,沃利斯构造了下面这些比值的表,表中对于 $p=0$,他简单地采用了 $y=1^n$ 下的面积:

$p\setminus n$	0	1	2	3	4	5	6	7	⋯
0	1	1	1	1	1	1	1	1	⋯
1	1	2	3	4	5	6	7	8	⋯
2	1	3	6	10	15	21	28	36	⋯
3	1	4	10	20	35	56	84	120	⋯
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋯

沃利斯在他的表中清楚地识别出了帕斯卡三角形. 他希望能插入对应于 $p = 1/2$, $p = 3/2, \cdots$ 的行以及对应于 $n = 1/2, n = 3/2, \cdots$ 的列, 从中他可以找到需要的值, 他将该值表示为□, 此时的两个参数都等于 1/2. 根据他对帕斯卡三角形的知识, 沃利斯意识到在他的表中关系式 $a_{p,n} = \frac{p+n}{n} a_{p,n-1}$ 成立, 式中 $a_{p,n} = \binom{p+n}{n}$ 指定 p 行 n 列的项. 对行 $p = 1/2$ 应用同样的规则, 他首先注意到 $a_{1/2,0} = 1$, 因为所有其它 0 列里的项都等于 1. 由此推出

$$a_{1/2,1} = \left(\frac{1/2+1}{1}\right) \cdot 1 = 3/2,$$

$$a_{1/2,2} = \left(\frac{1/2+2}{2}\right) \cdot \frac{3}{2} = \frac{5}{4} \cdot \frac{3}{2} = \frac{15}{8},$$

$$a_{1/2,3} = \frac{7}{6} \cdot \frac{5}{4} \cdot \frac{3}{2} = \frac{105}{48}, \cdots.$$

同样, 因为 $a_{1/2,1/2} = \square$, 他得到

$$a_{1/2,3/2} = \left(\frac{\frac{1}{2}+\frac{3}{2}}{3/2}\right)\square = \frac{4}{3}\square, a_{1/2,5/2} = \frac{6}{5} \cdot \frac{4}{3}\square, \cdots$$

于是行 $p = 1/2$ 为

$$1 \quad \square \quad \frac{3}{2} \quad \frac{4}{3}\square \quad \frac{15}{8} \quad \frac{8}{5}\square \quad \cdots$$

沃利斯能够类似地填出表中其余的部分, 但因为他关心的是计算 □, 他便考虑这一行中的比值. 因为相隔项的比值显然是不断减小的, 即对于所有的 k,

$$a_{1/2,k+2} : a_{1/2,k} > a_{1/2,k+4} : a_{1/2,k+2},$$

他假定这对相邻项也同样成立. 由此推出 $\square : 1 > 3/2 : \square$, 因此 $\square > \sqrt{3/2}$; 且 $3/2 : \square > (4/3)\square : 3/2$, 因此

$$\square < 3/2\sqrt{3/4} = [(3 \times 3)/(2 \times 4)]\sqrt{4/3},$$

并且, 同样的,

$$\square > [(3 \times 3)/(2 \times 4)]\sqrt{5/4}, \quad \square < [(3 \times 3 \times 5 \times 5)/(2 \times 4 \times 4 \times 6)]\sqrt{6/5}, \cdots$$

沃利斯由此能够断定 □(或 $4/\pi$) 可以用一个无穷积计算:

$$\square = \frac{4}{\pi} = \frac{3 \times 3 \times 5 \times 5 \times 7 \times 7 \times \cdots}{2 \times 4 \times 4 \times 6 \times 6 \times 8 \times \cdots}.$$

12.2.5 罗伯华和旋轮线

虽然一个无穷积也许不是沃利斯希望得到的那种面积结果, 同一时期的其他数学家在考虑幂曲线以外的曲线时也不得不满足于并非严格算术化的答案. 例如, 罗伯华在 1637 年前后确定了旋轮线 —— 由沿直线滚动的轮毂上附着的一点画出的曲线 —— 下的面积. 罗伯华将这种曲线定义为: "使圆 *AGB* 的直径 *AB* 沿切线 *AC* 移动, 并始终保持同原来的位置平行, 直到它到达位置 *CD*, 并令 *AC* 等于半圆 *AGB*(图 12.12). 同时, 使点 *A* 在半圆 *AGB* 以这种方式运动使得 *AB* 沿 *AC* 运动的速度可以等于点 *A* 沿半圆 *AGB* 运动的速度. 然后, 当 *AB* 到达位置 *CD* 时, 点 *A* 将达到位置 *D*. 点 *A* 由两种运动带动 —— 它本身沿半圆 *AGB* 的运动和直径沿 *AC* 的运动."[22]

罗伯华从把轴 *AC* 和半圆 *AGB* 分成无穷多相等的部分开始他的计算. 沿半圆的这些部分是 $AE = EF = FG = \cdots$, 而沿轴的部分是 $AM = MN = NO = \cdots$. 进而, 因为形成旋轮线的运动由沿半圆

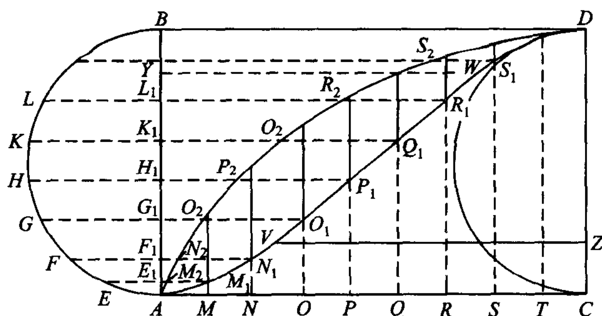


图 12.12 罗伯华确定旋轮线所围面积的方法.

和轴的等速运动合成,罗伯华令 $AE = AM, EF = MN \cdots$. 因为当直径的底在 M 时点 A 将在 E,与点 M 的水平距离与 E 到轴上的点 E_1 的水平距离相等的点 M_2 是旋轮线上一点. 同样与点 N 的水平距离与 F 到轴上的点 F_1 的水平距离相等的点 N_2 也是曲线上一点,图 12.12 标明的点 $O_2, P_2, \cdots$ 也一样. 罗伯华接着通过点 $M_1, N_1, O_1, \cdots$ 构造了一条新曲线,旋轮线的伴线,此处 M_1 的 x - 坐标与 M 相同,y - 坐标与 E 相同,以此类推. 用现代的概念说,该曲线由 $x(t) = at, y(t) = a(1-\cos t)$ 定义,或者用非参数形式表示,是 $y = a\left(1 - \cos\frac{x}{a}\right)$,此处 a 是圆的半径. 旋轮线本身由 $x(t) = a(t - \sin t), y(t) = a(1-\cos t)$ 所定义.

为确定旋轮线的半个拱下的面积,罗伯华首先确定出旋轮线和它的伴线间的面积等于生成圆面积的一半. 这可以从卡瓦列里原理推出,因为 $M_1M_2 = EE_1, N_1N_2 = FF_1, \cdots$ 并且对应的线对都在同样的高度. 为完成他的计算,罗伯华注意到对于区域 $ACDM_1$ 的每个线段 VZ,都在区域 AM_1DB 中对应着一个相等的线段 WY. 所以,还是根据卡瓦列里原理,旋轮线的伴线平分矩形 $ABCD$. 因为矩形的面积等于圆周长的一半与直径的乘积(或者 $2\pi a^2$),伴线下的面积等于生成圆的面积(πa^2). 由此得出,旋轮线的半个拱下的面积等于生成圆的面积的 3/2 或者说整个拱下的面积等于圆面积的三倍.

罗伯华的旋轮线的伴线实际上是一条余弦曲线,虽然他没有识别出这点. 但在同一部著作中,他的确画出并且可能是第一次画出了被确定为正弦曲线的曲线,虽然它只包含了四分之一圆的正弦. 并且,罗伯华能够确定出该曲线下的面积同定义该正弦的半径上的正方形相等. 大约 20 年后,帕斯卡在一篇题为《关于四分之一圆的正弦的论文》的短文中,能够找出该曲线的任何一部分下的面积. 考虑圆的四分之一 ABC 并令 D 是任意一点,从该点引正弦 DI 到半径 AC(图 12.13). 帕斯卡然后画了一个"小" 切线 EDE' 以及到半径的垂线 $ER, E'R'$. 他的论断是"任何四分之一圆弧的正弦的和等于两个端点正弦间部分的底与半径的乘积."[23] 帕斯卡用"正弦的和"表示由每个正弦与由切线 EE' 表示的无限小弧相乘组成的无限小矩形之和. 因此,帕斯卡的定理用现代术语表示就是

$$\int_{\alpha}^{\beta} r\sin\theta \, \mathrm{d}(r\theta) = r(r\cos\alpha - r\cos\beta).$$

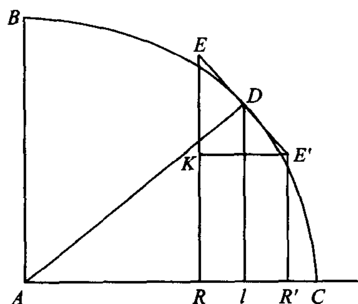


图 12.13 帕斯卡的正弦曲线下方的面积.

在证明中,帕斯卡注意到三角形 EKE' 和 DIA 是相似的. 所以 $DI : DA = E'K : EE' = RR' : EE'$,因此 $DI \cdot EE' = DA \cdot RR'$. 换言之,由正弦和无限小弧(或者切线)形成的矩形等于由半径和在弧的两端之间的轴的部分形成的矩形,或者 $r\sin\theta \mathrm{d}(r\theta) = r(r\cos(\theta + \mathrm{d}\theta) - r\cos(\theta)) = r(\mathrm{d}(r\cos\theta))$. 汇总两个给定角之

间的矩形就得到引用的结果.虽然该结果被证明是重要的,而且虽然帕斯卡立即将它推广,用来求得正弦的幂的积分公式,帕斯卡工作最重要的方面却是“微分三角形”EKE' 的出现.莱布尼茨对帕斯卡这项特殊工作的研究对他自己建立起在面积问题和切线问题间联系是至关紧要的.

12.2.6 等轴双曲线下的面积

我们关于 17 世纪中期对面积问题的解的最后一个例子是比利时数学家圣文森特的格利高里(Gregory of St. Vincent,1584—1667)在双曲线 $xy = 1$ 的工作(图 12.14).格利高里在他的《几何著作》中证明了,如果对 $i = 1,2,3,4$,(x_i, y_i) 是这个双曲线上的四个点,满足 $x_2 : x_1 = x_4 : x_3$,则在区间$[x_1, x_2]$上该双曲线下的面积等于在区间$[x_3, x_4]$上该双曲线下的面积(图 12.15).为证明这点,将区间$[x_1, x_2]$在点 $a_i, i = 0,\cdots,n$ 处分成子区间.因为 $x_2 : x_1 = x_4 : x_3$,由此得出 $x_3 : x_1 = x_4 : x_2 = v$ 或者 $x_3 = vx_1, x_4 = vx_2$.因此人们可以方便地将区间$[x_3, x_4]$在点 $b_i = va_i, i = 0,\cdots,n$ 分割.如果用矩形内接以及外接在区间$[a_j, a_{j+1}]$上和$[b_j, b_{j+1}]$上的双曲线下面积 A_j 和 B_j 周围,则可以直截了当地计算相应的不等式:



图 12.14 格利高里的《几何著作》扉页图.他宣称他已将圆化方.(来源:USMA 图书馆的特殊收藏部,西点,纽约.)

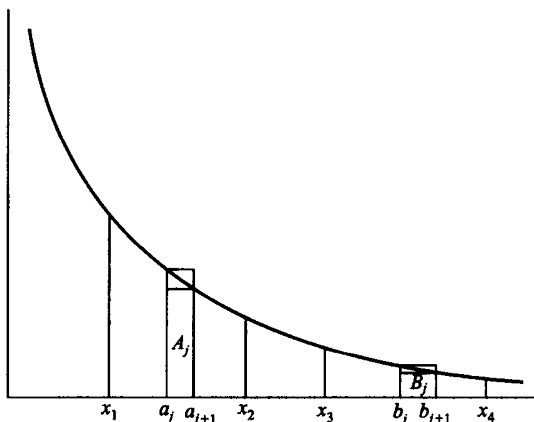


图 12.15　格利高里与双曲线 $xy = 1$ 下方的面积.

$$(a_{j+1} - a_j)\frac{1}{a_{j+1}} < A_j < (a_{j+1} - a_j)\frac{1}{a_j} \quad \text{和} \quad (b_{j+1} - b_j)\frac{1}{b_{j+1}} < B_j < (b_{j+1} - b_j)\frac{1}{b_j}.$$

将值 $b_j = va_j$ 代入第二组不等式,得

$$(a_{j+1} - a_j)\frac{1}{a_{j+1}} < B_j < (a_{j+1} - a_j)\frac{1}{a_j}.$$

因而两个双曲线下的区域被同样面积的矩形所夹挤.因为两个区间都可以分成任意小的子区间,由此得出两个双曲线下的面积相等.

在比利时的耶稣会教士阿尔方索·安东尼·德·撒拉萨(Alfonso Antonio de Sarasa,1618—1667)在 1649 年读格利高里的著作时,他即刻注意到这个计算意味着从 1 到 x 的双曲线下的面积 $A(x)$ 具有对数性质 $A(\alpha\beta) = A(\alpha) + A(\beta)$.(因为比值 $\beta : 1$ 等于比值 $\alpha\beta : \alpha$,从 1 到 β 的面积等于从 α 到 $\alpha\beta$ 的面积.因为从 1 到 $\alpha\beta$ 的面积是从 1 到 α 和从 α 到 $\alpha\beta$ 的面积之和,立即得到对数性质.)因此如果人们计算部分双曲线 $xy = 1$ 下的面积,就可以计算对数.对计算这些面积的方法的寻求在 17 世纪 60 年代引出了牛顿和其他人的幂级数方法,这些方法在牛顿版的微积分中是至关重要的.

12.3　幂级数

1668 年,尼古劳斯·梅卡托(Nicolaus Mercator,1620—1687)发表了他的《对数课程》,中间出现了对数的幂级数展开式.梅卡托读了撒拉萨对对数同双曲线下的面积相联系的暗示,从沃利斯那里学了如何计算无限幂和的某些比,决定利用这些无限和来计算对数 $\log(1 + x)$(从 0 到 x 的双曲线 $y = 1/(1 + x)$ 下的面积 A).他将区间 $[0, x]$ 分为长度为 x/n 的 n 个子区间并用以下和式逼近 A:

$$\frac{x}{n} + \frac{x}{n}\left(\frac{1}{1 + \frac{x}{n}}\right) + \frac{x}{n}\left(\frac{1}{1 + \frac{2x}{n}}\right) + \cdots + \frac{x}{n}\left(\frac{1}{1 + \frac{(n-1)x}{n}}\right).$$

因为每一项 $\frac{1}{1 + (kx/n)}$ 是几何数列 $\sum_{j=0}^{\infty}(-1)^j\left(\frac{kx}{n}\right)^j$,由此可得

$$\begin{aligned} A &\approx \frac{x}{n} + \frac{x}{n}\sum_{j=0}^{\infty}(-1)^j\left(\frac{x}{n}\right)^j + \frac{x}{n}\sum_{j=0}^{\infty}(-1)^j\left(\frac{2x}{n}\right)^j + \cdots + \frac{x}{n}\sum_{j=0}^{\infty}(-1)^j\left(\frac{(n-1)x}{n}\right)^j \\ &= n\frac{x}{n} - \frac{x^2}{n^2}\sum_{i=1}^{n-1} i + \frac{x^3}{n^3}\sum_{i=1}^{n-1} i^2 + \cdots + (-1)^j\frac{x^{j+1}}{n^{j+1}}\sum_{i=1}^{n-1} i^j + \cdots \end{aligned}$$

$$= x - \frac{\sum_{i=1}^{n-1} i}{n \cdot n} x^2 + \frac{\sum_{i=1}^{n-1} i^2}{n \cdot n^2} x^3 + \cdots + (-1)^j \frac{\sum_{i=1}^{n-1} i^j}{n \cdot n^j} x^{j+1} + \cdots.$$

根据沃利斯的结果,如果 n 为无穷大,该表达式中 x^{k+1} 的系数将等于$\frac{1}{k+1}$.因此

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots,$$

这是一个能使人们轻易计算对数的实际值的 x 的幂级数.

其它超越函数的幂级数是 1670 年前后在苏格兰由詹姆斯·格利高里(James Gregory, 1638—1675)发现并通报给皇家学会秘书约翰·柯林斯(John Collins,1625—1683)的,通报中没有任何如何发现这些级数的线索.例如,在 1670 年 12 月 19 日的一封信里,格利高里写下,正弦为 B 的弧(此处圆的半径为 R)可以表示为

$$B + \frac{B^3}{6R^2} + \frac{3B^5}{40R^4} + \frac{5B^7}{112R^6} + \frac{35B^9}{1152R^8} + \cdots.$$ [24]

用现代的术语说,格利高里的级数是$\frac{1}{R}\arcsin\frac{B}{R}$ 的级数,当 $R=1$ 时,它可以写成

$$\arcsin x = x + \frac{x^3}{6} + \frac{3x^5}{40} + \frac{5x^7}{112} + \frac{35x^9}{1152} + \cdots.$$

同样地,格利高里在 1671 年 2 月 15 日的信里包含但不限于给定正切 x 求弧长 y 或者相反问题的数列,用现代符号可以写为

$$\arcsin x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots,$$

$$\tan y = y + \frac{y^3}{3} + \frac{2y^5}{15} + \frac{17y^{17}}{315} + \frac{3233y^9}{181\,440} + \cdots.$$ [25]

对在格利高里收到的信的边角和其它空白处发现的他的注释的研究使现代学者们相信,格利高里使用了 40 多年后才首次发表的泰勒级数的法则的一种,用现代的符号,该法则将一个函数表达为它的导数的级数:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots.$$

不论格利高里是如何推出这些级数的,人们发现由牛顿在 17 世纪 60 年代中期发现的反正切级数以及正弦和余弦的级数都在大约更早的 200 年前在南印度被发现.这些级数以梵语诗歌的形式出现在 Tantrasangraha-vyākhyā(约 1530 年),这是一个对大约三十年前喀拉拉·伽吉亚·尼拉康达(Kerala Gargya Nīlakantha 1445—1545)的一部著作的评论.令人惊讶的是,在一部以印度西南喀拉拉邦的马拉雅拉姆语写成的著作 Yuktibhasa 中存在着这些级数详细的推导,这部著作是由加斯特德维(Jyesthadeva 1530—1610)写的,他把反正切级数归功于住在克钦附近的更早的数学家马德哈拉(Madhava,1340—1425).

那段给出反正切级数的梵语诗歌也许可以翻译如下:

> 给定正弦与半径的乘积被余弦除,是第一个结果.从第一[以及第二、第三等]结果,通过不断乘以正弦的平方除以余弦的平方[依次]可以得到一串结果.用以上结果依次除以奇数 1,3 等[得到完整的项的序列].从奇数项的和中减去偶数项的和.结果就是弧.在这一场合下……弧的正弦或者补弧的正弦不论多么小在这里都应该被取为[给定的正弦];否则用[以上]重复步骤得到的项就不会趋于 0.[26]

注意到,作者意识到了只有当 $x=\tan y\leqslant 1$ 时级数才收敛,不难将这段话翻译成格利高里发现的用现代符号表示的同一个反正切级数.

加斯特德维对反正切级数的推导是从下面一条引理开始的,为简单起见,圆的半径被取为 1.

引理　设 BC 是圆心在 O 的圆的一小段弧.如果 OB,OC 分别交从圆上任意一点 A 引出的切线于 B_1,C_1,则弧 BC 由下式近似给出:$\text{arc }BC\approx B_1C_1/(1+AB_1{}^2)$(图 12.16).

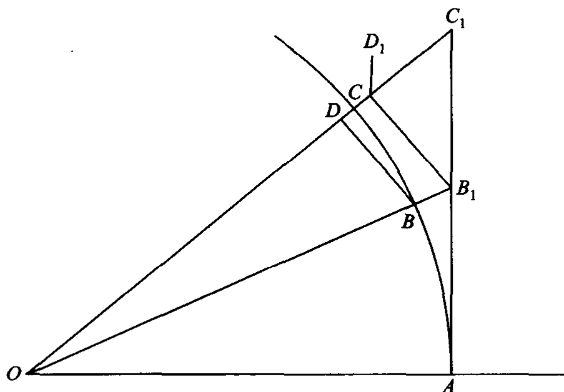


图 12.16　加斯特德维对反正切级数的推导.

如果向 OC 引垂线 BD,B_1D_1,根据相似性可得:

$$\frac{BD}{B_1D_1}=\frac{OB}{OB_1}=\frac{1}{OB_1}\quad \text{和}\quad \frac{B_1D_1}{B_1C_1}=\frac{OA}{OC_1}=\frac{1}{OC_1}.$$

因此 $BD=\dfrac{B_1C_1}{OB_1\cdot OC_1}$.当弧 BC 很小时,$OB_1\approx OC_1$ 因此

$$BC\approx BD=\frac{B_1C_1}{OB_1^2}=\frac{B_1C_1}{1+AB_1^2}.$$

将弧 AC 对应的正切 $t=AC_1$ 分成 n 个相等部分,对每部分应用引理,然后令 n 无限变大得

$$\begin{aligned}\arctan t&=\lim_{n\to\infty}\sum_{r=0}^{n-1}\frac{t/n}{1+\left(\frac{rt}{n}\right)^2}\\&=\lim_{n\to\infty}\sum_{r=0}^{n-1}\frac{t}{n}\left[1-\left(\frac{rt}{n}\right)^2+\left(\frac{rt}{n}\right)^4-\cdots+(-1)^k\left(\frac{rt}{n}\right)^{2k}+\cdots\right]\\&=\lim_{n\to\infty}\left[\frac{t}{n}+\frac{t}{n}\left(1-\frac{t^2}{n^2}+\frac{r^4}{n^4}-\cdots\right)+\frac{t}{n}\left(1-\frac{2^2t^2}{n^2}+\frac{2^4t^4}{n^4}-\cdots\right)\right.\\&\quad\left.+\frac{t}{n}\left(1-\frac{3^2t^2}{n^2}+\frac{3^4t^4}{n^4}-\cdots\right)+\cdots+\frac{t}{n}\left(1-\frac{(n-1)^2t^2}{n^2}+\frac{(n-1)^4t^4}{n^4}-\cdots\right)\right]\\&=\lim_{n\to\infty}\left[t-\frac{t^3}{n^3}(1^2+2^2+\cdots+(n-1)^2)+\frac{t^5}{n^5}(1^4+2^4+\cdots+(n-1)^4)-\cdots\right].\end{aligned}$$

为完成推导,加斯特德维需要证明累积的幂的和可以被更简单的表达式取代,特别是沃利斯定理成立:

$$\lim_{n\to\infty}\frac{1}{n^{p+1}}\sum_{j=1}^{n-1}j^p=\frac{1}{p+1}.$$

虽然他普遍地陈述了该结果,但他不能对所有的 p 推导出它.因此,尽管我们将用现代通用的符号给出他的结果,加斯特德维只对小 p 值的情况作了验证并假定它们对所有的值成立.但他的确从沃

利斯的定理在 $p = 1$ 时成立这个简单的事实开始使用了归纳论证.他从海塞姆公式7.2的一种形式开始：

$$n\sum_{j=1}^{n} j^{p-1} = \sum_{j=1}^{n} j^{p} + \sum_{k=1}^{n-1}\sum_{j=1}^{k} j^{p-1}.$$

接着,他证明了如果沃利斯的结果对指数 $p-1$ 成立,则当 n 变大时：

$$\sum_{k=1}^{n-1}\sum_{j=1}^{k} j^{p-1} \approx \frac{1}{p}\sum_{j=1}^{n} j^{p}.$$

由此可得,当 n 变大时,

$$n\sum_{j=1}^{n-1} j^{p-1} \approx \left(1+\frac{1}{p}\right)\sum_{j=1}^{n-1} j^{p} = \frac{p+1}{p}\sum_{j=1}^{n-1} j^{p},$$

因此

$$\lim_{n\to\infty}\frac{1}{n^{p+1}}\sum_{j=1}^{n-1} j^{p} = \lim_{n\to\infty}\frac{1}{n^{p+1}}\frac{p}{p+1}n\frac{n^{p}}{p} = \frac{1}{p+1}.$$

所以根据归纳法,沃利斯的定理成立,而我们可以利用它替换 arctan t 表达式中的幂和.最终结果是

$$\arctan t = t - \frac{t^3}{3} + \frac{t^5}{5} - \frac{t^7}{7} + \cdots.$$

为什么印度的作者对该数列有兴趣?他们主要的目的似乎是计算圆弧的长度,此值对天文上的目的是必需的.该数列使这种计算成为可能.例如,直接代入 $t = 1$ 得到

$$\pi/4 = 1 - 1/3 + 1/5 - 1/7 + \cdots.$$

但因为这一数列收敛得很慢,有必要作各种各样的改进.例如,Tantrasangraha-vyākhyā 一书中就包含了其它收敛快得多的数列,包括

$$\frac{\pi}{4} = \frac{3}{4} + \frac{1}{3^3-3} - \frac{1}{5^3-5} + \frac{1}{7^3-7} - \cdots.$$

非常有趣的是,使欧洲的作者们认识到切线问题同面积问题有关的是确定曲线弧长的同一个问题.

12.4　曲线求长法和基本定理

笛卡儿在他的《几何学》中声称人类的才智无法发现确定曲线和直线比值的严格而精确的方法.仅仅在笛卡儿写下这些话的二十年后,就有人类的才智证明他是错误的.第一个曲线长度的确定可能是英国人威廉·耐尔(William Neile,1632—1670)在沃利斯的建议下于1657年确定了半三次抛物线 $y^2 = x^3$ 的长度.接下来的两年里圣保罗教堂和伦敦许多其它教堂的建筑师克里斯托弗·雷恩(Christopher Wren,1632—1723)确定了旋轮线的长度,惠更斯将求抛物线长度的问题化为了求双曲线下面积的问题.但最普遍的过程是由亨德里克·范·休莱特(Hendrick van Heuraet,1634—1660(?))发现,它出现在范·舒滕编的1659年拉丁版的笛卡儿的《几何学》中.

12.4.1　范·休莱特的工作

范·休莱特在其论文《论曲线到直线的转换》的开始证明了作长度等于给定弧的直线段的问题等价于找某一曲线下的面积的问题.设 P 为曲线 α 的弧 MN 上的任意一点(图12.17).从 P 到轴的法线 PS 的长度可以用笛卡儿的方法确定.取任意线段 σ,范·休莱特通过比例 $P'R : \sigma = PS : PR$ 定

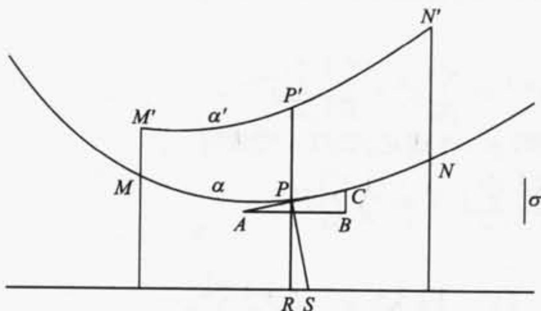


图 12.17　范 · 休莱特的求曲线弧长法.

义一条新曲线 α',其中 P' 是同 P 相伴的 α' 上的一点.(引入 σ 是为了使所有的比都是线段的比.)作微分三角形 ACB,其中 AC 是点 P 处 α 的切线,他注意到 $PS : PR = AC : AB$.用现代的符号表示,如果 $AC = \mathrm{d}s$ 而 $AB = \mathrm{d}x$,则范 · 休莱特的比得出 $P'R : \sigma = \mathrm{d}s : \mathrm{d}x$ 或者 $\sigma\mathrm{d}s = P'R\mathrm{d}x$.因为 MN 的长度由曲线 MN 上无限小切线的和,或者等价地,弧的无限小的小段的和给出,范 · 休莱特总结出 $\sigma \cdot$(MN 的长度) $= M'$ 和 N' 之间的曲线 α' 下的面积.因此,如果从 α 的方程能够推出 α' 的方程并计算出它下面的面积,MN 的长度就也能被计算出.再使用现代的术语,有

$$z = P'R = \sigma\frac{\mathrm{d}s}{\mathrm{d}x} = \sigma\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2},$$

范 · 休莱特的过程可以写为

$$\sigma \cdot (MN\text{ 之长}) = \int_a^b z\mathrm{d}x = \int_a^b \sigma\sqrt{1+\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2}\mathrm{d}x,$$

其中 a,b 表示 M 和 N 的 x 坐标,它本质上是现代的弧长公式.

人物小传	亨德里克 · 范 · 休莱特(1634—1660)(Hendrick Van Heuraet)
	范 · 休莱特生于荷兰的哈勒姆,1653 年到莱顿跟随范 · 舒滕学习数学.[27] 他从前一年去世的布商父亲那里继承的遗产使他相当富裕,可以研究和旅游而不用担心物质资助.他的早期数学工作显露出他远大的前程,范 · 舒滕不仅发表了他关于求弧长的论文还发表了他在拐点方面的工作.但据我们所知,范 · 休莱特英年早逝.他在 1660 年初以后的活动就已无据可考.

范 · 休莱特用了相伴曲线下的面积能被实际计算出的少数曲线中的一个说明了他的过程,这就是早些时候耐尔考虑过的半立方抛物线 $y^2 = x^3$.利用笛卡儿的法线法,他计算了必须有重根的方程是 $x^3 + x^2 - 2vx + v^2 - n^2 = 0$.利用找重根的胡德法则,他给这个方程的各项乘以 3,2,1,0 得到 $3x^3 + 2x^2 - 2vx = 0$.因此 $v - x = 3x^2/2$ 并且 $PS = \sqrt{\frac{9}{4}x^4 + x^3}$.令 $\sigma = 1/3$,范 · 休莱特取

$$z = P'R = \sigma \cdot \frac{PS}{PR} = \frac{1}{3}\,\frac{\sqrt{(9/4)x^4 + x^3}}{\sqrt{x^3}} = \sqrt{\frac{1}{4}x + \frac{1}{9}},$$

或者,等价地,$z^2 = (1/4)x + 1/9$,定义了新的曲线 α'.范 · 休莱特轻易地辨认出该曲线是抛物线,他知道如何计算抛物线下的面积.从 $x = 0$ 到 $x = b$ 的半立方抛物线的长度这时等于这一面积除以 σ,即

$$\sqrt{\left(b+\frac{4}{9}\right)^3}-\frac{8}{27}.$$

范·休莱特评论说人们可以类似地明确确定出曲线 $y^4 = x^5, y^6 = x^7, y^8 = x^9, \cdots$ 的长度,然后他用求抛物线 $y = x^2$ 的一段弧长这个更困难的问题结束了论文,该问题取决于确定双曲线 $z = \sqrt{4x^2+1}$ 下的面积.这个问题在 1659 年还没有得到满意解决.无论如何,范·休莱特的方法很快便广为人知.特别是,微分三角形和与给定曲线相伴将新曲线的运用有助于其他人获得将切线问题同面积问题相联系的主意.

12.4.2　格利高里和基本定理

在将切线问题同面积问题相联系的数学家中有伊萨克·巴罗(Isaac Barrow,1630—1677)和詹姆斯·格利高里,两个人都决定把他们在游历法国、意大利和荷兰时搜集的与切线、面积和求弧长的资料整理起来加以系统的表述.那么毫不奇怪,巴罗的《几何讲义》(1670)和格利高里的《几何的通用部分》(1668)包含着许多用相似方法表述的相同材料.实际上,这些著作都是论述今天属于微积分领域的材料的专著,但表述采用了两位作者在他们大学时代学到的几何风格.两人都没能将材料翻译成对解决问题有用的计算方法.

人物小传	詹姆斯·格利高里(1638—1675)(James Gregory)
	格利高里在阿伯丁的马里沙尔学院毕业后,在 1663 年离开了苏格兰并在国外度过了 5 年光阴,在意大利的帕杜阿跟随托利拆里的学生斯蒂方诺·德格里·安格里学习.在那里他写下了他的前两部数学著作.1668 年,他回国担任圣安德鲁斯的数学教授,他在那里花费了大量时间讲授初等数学.他同在伦敦的约翰·柯林斯的通信是他同外部数学界的惟一联系.1673 年,他因政治问题被迫离开圣安德鲁斯,但不久后便担任了爱丁堡的教授.不幸的是,他在 1675 年 10 月因中风而失明,不久便与世长辞.

作为一个例子,我们来考虑格利高里是如何表述微积分基本定理的,这一结果把面积和切线的思想联系在了一起.该结果是格利高里在研究他从范·休莱特著作中找到的弧线长度的一般问题时得到的自然成果.考虑一个单调递增的曲线 $y = y(x)$,有两条另外的曲线同它相关联,法曲线 $n(x) = y\sqrt{1+(\mathrm{d}y/\mathrm{d}x)^2}$ 和 $u(x) = cn/y = c\sqrt{1+(\mathrm{d}y/\mathrm{d}x)^2}$,此处 c 是一给定常数.现在在给定点构造微分三角形 $\mathrm{d}x, \mathrm{d}y, \mathrm{d}s$,它与由纵坐标 y,次法线 v 和法线 n 组成三角形相似,由此他论证出 $y : n = \mathrm{d}x : \mathrm{d}s = c : u$,因此 $u\mathrm{d}x = c\mathrm{d}s$ 且 $n\mathrm{d}x = y\mathrm{d}s$(图 12.18).像范·休莱特一样,将第一个方程在曲线上求和,格利高里证明出弧长 $\int \mathrm{d}s$ 可以用曲线 $\frac{1}{c}u(x)$ 下的面积表示.第二个方程的和使得格利高里可以证明 $n = n(x)$ 下的面积同绕 x - 轴旋转原曲线形成的曲面的面积只差一个常数因子.格利高里利用内接和外接矩形、双归谬法通过仔细的阿基米德式的论证证明了这两个结果.

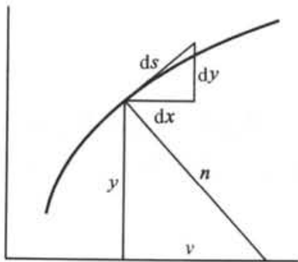


图 12.18　格里高利的微分三角形.

在证明了弧长可以通过面积得到之后,格利高里通过提出逆问题取得了根本性的进展.人们能否找到一个弧长 s 同给定曲线 $y(x)$ 下的面积成常数比的曲线 $u(x)$?用现代的符号表示,格利高里要求确定 u,使得

$$c\int_0^x \sqrt{1+\left(\frac{\mathrm{d}u}{\mathrm{d}x}\right)^2}\mathrm{d}x = \int_0^x y\mathrm{d}x.$$

但这意味着 $c^2(1+(\mathrm{d}u/\mathrm{d}x)^2) = y^2$ 或者 $\mathrm{d}u/\mathrm{d}x = (1/c)\sqrt{y^2-c^2}$. 换言之,格利高里必须确定一条曲线,其切线斜率等于给定的函数. 令 $z=\sqrt{y^2-c^2}$,格利高里简单地定义 $u(x)$ 是从原点到 x 的曲线 z/c 下的面积. 他的任务此时成了证明该曲线的切线的斜率由 z/c 给出. 他实际证明的是连接曲线 u 上一点 K 与同 K 的 x - 坐标相距为 cu/z 的轴上一点的直线与曲线相切于 K.

格利高里关键的进步则是将两个给定的 x 值间特定曲线下面积的思想抽象为作为变量函数的面积的思想. 换言之,他构造了任意点 x 上的纵坐标等于从给定点到 x 的原曲线下的面积的新曲线. 一旦这一思想被提出,作新曲线的切线并证明它在点 x 处的斜率始终等于该点处原来的纵坐标实际上就不是什么难事.

12.4.3 巴罗和基本定理

格利高里仅在处理弧长时才有作新曲线的思想,另一方面,伊萨克 · 巴罗则以更为普遍的形式表述了基本定理的一部分作为他《几何讲义》第十讲的第 11 个命题.

定理 设 ZGE 为任意一条曲线,它的轴是 AD,所有的纵坐标都相对于该轴. AZ, PG, DE 从初始的纵坐标 AZ 连续增加. 再设 AIF 是这样一条曲线,如果垂直于 AD 作任意直线 EDF 交曲线于 E, F, 交 AD 于 D, 则有由 DF 和给定长度 R 围成的矩形等于截取的面积 $ADEZ$. 又设 $DE : DF = R : DT$ 并连接 FT, 则 TF 与 AIF 相切(图 12.19). [28]

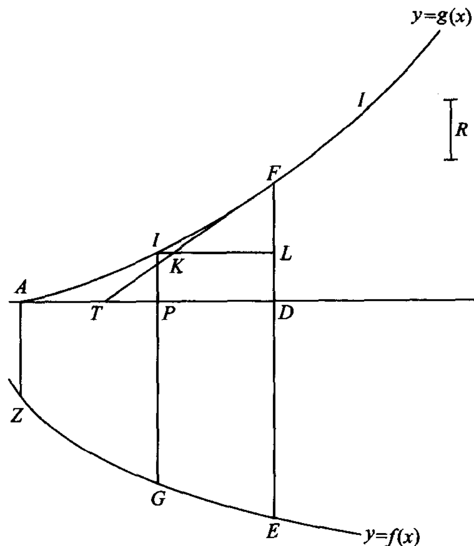


图 12.19 基本定理的巴罗形式.

像格利高里一样,巴罗从一条曲线 ZGE 开始,我们将其写作 $y = f(x)$, 作新曲线 $AIF = (g(x))$ 使得 $Rg(x)$ 始终等于 $f(x)$ 在一定点和变量 x 间围成的面积. 用现代的符号,

$$Rg(x) = \int_a^x f(x)\mathrm{d}x.$$

巴罗然后证明了 $g(x)$ 的次切距 $t(x)$ 由 $Rg(x)/f(x)$ 给出,或者

$$g'(x) = \frac{g(x)}{t(x)} = \frac{f(x)}{R} \quad 或 \quad \frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(x)\mathrm{d}x = f(x).$$

巴罗通过说明直线 TF 总是位于曲线之外证明了这一结果.如果 I 是曲线 $g(x)$ 上 F 向 A 的一侧的任意一点,并且如果作 IG 平行于 AZ 以及 KL 平行于 AD,曲线的性质说明 $LF : LK = DF : DT = DE : R$ 或者 $R \cdot LF = LK \cdot DE$.因为 $R \cdot IP$ 等于 $APZG$ 的面积,由此得出 $R \cdot LF$ 等于 $PDEG$ 的面积.所以 $LK \cdot DE = PDEG$ 的面积 $< PD \cdot DE$.因此 $LK < PD$ 或者说 $LK < LI$,而切线在点 I 位于曲线以下.同样的论证适用于位于 F 远离 A 的一侧的点 I.

人物传记	伊萨克·巴罗(1630—1677)(Isaac Barrow)
	巴罗 1643 年进入剑桥大学三一学院学习,1648 年取得学士学位,1652 年取得硕士学位.因为他对皇家的同情,在 1655 年被学校罢黜而被禁止取得教授席位.他利用这一机会在大陆游历了 4 年并在法国、意大利和荷兰学习数学.在复位时期,他回到剑桥,接受圣职,并成为希腊语的钦定讲座教授.1662 年,他兼任伦敦的格雷斯海姆几何教授,第二年成为剑桥大学的首任卢卡斯数学教授.随后的几年里,他在初等数学、几何和光学方面讲授了几门课程,1669 年他辞职作了伦敦的皇家牧师.1673 年,他回到三一学院任院长,两年后成为大学副校长,但他在 1677 年离世,原因可能是过量服用了药品.

在第十一讲的第 19 个命题中,巴罗通过在曲线 $Rf'(x)$ 下的区域里的无限小矩形和在大的矩形 $R(f(b) - f(a))$ 中的无限小矩形间建立对应,证明了基本定理的第二部分,即

$$\int_a^b Rf'(x)\mathrm{d}x = R(f(b) - f(a)),$$

巴罗是如何发现切线和面积问题的互逆关系的?巴罗没有明确地告诉我们,但对《几何讲义》早期部分的仔细阅读显示出他经常把曲线看作是由运动的动点生成的.例如他说明了这样产生的曲线在点 P 处切线的斜率等于该动点在点 P 处的速度.并且,他还用曲线变动的坐标表示点的变动的速度,曲线的轴代表时间."因此,如果通过代表时间的所有点作处理得互不重合的直线(即平行线),作为这些平行线——每条都表示了与作出这些平行线的点相对应的速度——的集合得到的平面准确地对应着速度的集合,因而也可以最方便地修改来表示穿过的距离."[29] 用速度曲线下的面积表示距离的主意归于伽利略和奥尔斯姆,但是这一主意同作为切线斜率的速度的概念相结合,便能使巴罗轻易地理解到微分和积分过程的互逆关系.

巴罗在发表他的《几何讲义》紧前面的年代里是剑桥大学的卢卡斯数学教授.不知道伊萨克·牛顿是否听过巴罗的任何讲座,但他很可能受到过巴罗的运动生成曲线的思想的影响.事实上牛顿对巴罗的书提出过若干改进,特别是他建议巴罗收入一种以微分三角形为基础的计算切线的代数方法.该方法包括在给定曲线上的点 M 处作微分三角形 NMR 并利用无限小三角形中对应的 $MR = a$ 与 $NR = e$ 的比计算 $MP = y$ 与 $PT = t$ 的比(图 12.20).例如,如果曲线是 $y^2 = x^3$,巴罗用 $y + a$ 替代 y,用 $x + e$ 替代 x,得到 $(y + a)^2 = (x + e)^3$ 或者 $y^2 + 2ay + a^2 = x^3 + 3x^2e + 3xe^2 + e^3$.接着他剔除掉所有包含 a 或者 e 的幂或者两个的乘积的项,"因为这些项没有价值,"得到 $y^2 + 2ay = x^3 + 3x^2e$.接着,"丢弃所有含有表示已知或确定数值的字母的项……因为这些项移到方程的一边,总是等于 0,"他还剩下 $2ay = 3x^2e$.最后一步,他用 a 替代 y,用 e 替代 t 得到比值 $y : t$.在此例中,结果是 $y : t = 3x^2 : 2y$.巴罗进一步注意到"如果计算中有任何曲线的无穷小的弧,可以用无穷小的一段切线或者任何等价于它的直线替代它."[30] 巴罗因而通过完全忽略掉等同的概念改进了费马的方法.他除了提到他经常在自己的计算中使用该方法外并没有试图说明他的方法的合理性

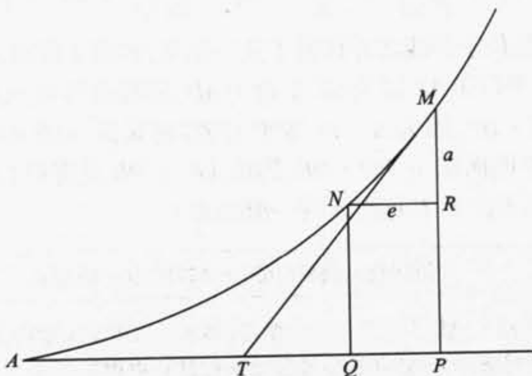


图 12.20　巴罗的微分三角形.

(补遗 12.2).

补遗 12.2	巴罗发明了微积分吗?
	即使承认巴罗知道计算切线和面积的代数过程并且也了解基本定理,他是否应该被当作是微积分的一个发明人.答案肯定是否定的.巴罗的所有工作都是用传统的几何方式表述的.他似乎并不了解他的课本中表述的两个定理的基本性质.巴罗没有说过它们特别重要.它们被表述为许多有关切线和面积的几何结果中的两个.并且巴罗从未利用它们计算过面积.也许如果没有牛顿跟上来的话,巴罗会看到这些定理可能被派什么用场.但因为他意识到牛顿的才能在他之上,并且因为他更关心探求神学方面的兴趣,巴罗放弃了对数学的研究,把它和微积分的发明留给了他年轻的同事.

巴罗和格利高里的工作可以被认为是所有 17 世纪面积和切线计算方法的一个高峰,但他们没有人在 1670 年能将这些方法融合成一种真正的计算和解决问题的工具.但在这一日期的 5 年前,在剑桥的斗室中孤军奋战的伊萨克·牛顿已经在用他强大的全神贯注的能力将他所有前辈们的工作整合扩展成为我们今天称为微积分的学科.

12.5　伊萨克·牛顿

据传记作家理查德·威斯法说,伊萨克·牛顿是"塑造了人类才智诸领域的寥寥无几的超级天才之一,一个无法归结为我们用以理解同类的标准的人."[31] 因为微积分仅仅是他对我们理解周围世界作出重大贡献的许多领域中的一个,因为新近由德里克·怀特赛德编辑的他的数学论文集填满了八大卷,对为什么他被认为是这样的"超级天才"的理由,我们在这里只能作简短一瞥.但是在后面几页将会搞清楚的是,在 17 世纪 60 年代的短短几年里牛顿成功地将他 17 世纪的前辈们发展出的关于切线和面积的所有材料统一并推广成为在我们今天 1000 页的微积分教科书中展示的神奇的解决问题的工具.

人物小传

伊萨克·牛顿(1642—1727)(Issac Newton)

牛顿1642年12月25日生于格兰瑟姆附近的沃尔索普村,此处在伦敦以北约100英里的地方.他的母亲在生他当年的10月就已守寡.在他3岁时,他母亲再嫁,把年幼的伊萨克留给他祖母照管.直到1653年,她在她第二个丈夫死后才重新回到沃尔索普.1655年,牛顿被送到格兰瑟姆就读地方文法中学.在这里,他掌握了拉丁语,这是传统中学课程的主干.并且,有点非同寻常的校长亨利·斯托克斯还引导他学习数学.牛顿不仅学习了基本的算术,还学习了像平面三角学和几何作图这样的高级科目.这使得他在1661年进入剑桥大学三一学院时就遥遥领先于他的同学.

但数学通常不是剑桥大学课程学习的一部分,甚至在巴罗1663年被任命为卢卡斯数学教授后依然如此.实际上,大学几乎压根就没有什么必修课.如果学生能住校4年并支付学费,就能取得学士学位.另一方面,因为1663年牛顿开始独立探索他在中学时代接触过的数学,大学对他研究什么并不特别关心倒是一件好事.他掌握了欧几里得几何以便理解三角学,然后是威廉·沃特莱德(William Oughtred,1574—1660)的《数学入门》,这是一本包含了算术和几何基本内容的通俗读物,接着又读了范·舒滕拉丁版的笛卡儿的《几何学》以及它数百页的注释,还读了韦达选集,最后还有沃利斯的《无穷算术》.因为伊萨克·巴罗在1664年开始了他关于数学基础的最早的一系列卢卡斯讲座,年长的数学家很可能鼓励过年轻的数学家,甚至可能从他的私人藏书中借给后者.但为了能全身心投入研究,牛顿需要大学经济资助的保障.这是通过1664年的奖学金,1667年的研究奖学金和1669年卢卡斯教授的任命来保证的,这一切可能都是通过了巴罗的影响力.

很明显,牛顿能在微积分的创立以及光学和力学基本原理的建立方面取得成功的一个主要原因是他高度的聚精会神的能力.如约翰·梅纳德·凯恩斯所写的:"我相信他智力的线索在于他连续全神贯注的沉思的超常的能力……他的特出天赋就是在头脑中一直保留一个纯智力问题直到他看透它为止的能力……我相信牛顿能够在头脑中把一个问题保留几个小时、几天甚至几个星期直到问题的秘密被他揭穿."[32] 牛顿聚精会神的能力被许多关于他的类似于有关阿基米德的故事所例证.例如,"当他在房间中有朋友要招待时,如果他走到书房去取一瓶酒,而他脑海中闪过一个主意,他会坐下来书写,全然忘记他的朋友."[33] 实际上,"考虑到所有那些没有用在研究上的失去的时间,……他很少离开房间,除非是在学期里,他在学校里作为卢卡斯教授授课."但当他讲课时,"很少的人去听他讲,更少的人能听懂,经常因为缺乏听众,他简直是对着空墙在讲."[34] 也许牛顿作为教授并不成功,但作为科学革命中的中心人物,他的著作对我们的生活持续不断地施加着影响.

牛顿在通过自学掌握了17世纪数学的全部成就后,从1664年后期到1666年后期花费了两年时间理出了他关于微积分的基本思想,工作的地点有时是在剑桥的房间里,有时是在沃尔索普的家中.虽然他的数学研究在他投身于包括光学、力学和炼金术在内的课题时会有所中断,他在随后的几年里作了更多的工作.至少有三回,牛顿把他的研究写成了适合发表的形式.不幸的是,因各种原因,牛顿从未发表三篇关于微积分的论文中的任何一篇.尽管如此,人们所说的《1666年10月流数简论》、1669年的《有无限多项方程的分析学》(后面简称为《分析学》)以及1671年的《论级数方法和流数法》(后面简称为《论方法》)都在一定程度上以手稿的方式在英国数学界传播并显示出牛顿新方法的巨大威力.因为后一篇论文概括并深化了两



图 12.21 苏联邮票上的牛顿.

篇较早论文的成果,我们将用它作为框架来研究牛顿的微积分,只在必要时参考另外两篇.

12.5.1 幂级数

1671 年的论文是从幂级数开始的.牛顿的中心思想是算术中的无穷小数和我们称为幂级数的无穷次"多项式"间的类比:

> 因为计算数和计算变量的运算非常相似…… 我很惊奇竟然没有人(如果你排除掉为双曲线求积的 N·梅卡托)把近来发展起来的用于小数的学说以相似的方式用到变量上,特别是因为这条路线在当时有望获得更为惊人的结果.因为这一类学说对于代数同小数的学说对于普通算术有着同样的关系,如果读者对算术和代数两个领域都熟悉并欣赏小数和直到无穷的代数项间的联系的话,它的加法、减法、乘法、除法和开方运算都可以借鉴后者的运算…… 正如小数的优点是当所有的分数和根都化为它们的时候它们在某种程度上呈现整数的性质那样,无穷变量级数的优点是更复杂的项的种类(例如分母为复数的分数,复数的根以及隐性方程的根)可以化为简单项的类别:即化为有简单分子和分母的分数的无穷级数而没有困扰其它类别的不可克服的障碍.[35]

牛顿进而用实例说明了无穷的变量级数或者说幂级数的优点,他把幂级数简单地看成了可以像处理普通多项式那样来处理的广义多项式.例如,分数 $1/(1+x)$ 可以简单地用长除法计算 1 除以 $1+x$ 而写成级数:

$$1 - x + x^2 - x^3 + x^4 - x^5 + \cdots.$$

同样地,人们可以用标准的确定平方根的算术算法来计算多项式的根,将其表示为幂级数.对 $\sqrt{1+x^2}$ 应用该方法,牛顿轻易地计算出结果:

$$1 + \frac{x^2}{2} - \frac{x^4}{8} + \frac{x^6}{16} - \frac{5x^8}{128} + \frac{7x^{10}}{256} - \cdots.$$

"隐性方程"的化简,即把方程 $f(x,y)$ 中的 y 表示为 x 的幂级数,要更困难一些.牛顿相信,因为人们对求方程 $f(y)=0$ 的数值解的方法不完全熟悉.因此牛顿以 $y^3-2y-5=0$ 为例解释了他的解这类方程的方法.他注意到整数 2 可以作为对一个根的最初近似.接着他令 $y=2+p$ 并代入原方程得到新方程 $p^3+6p^2+10p-1=0$.因为 p 小,牛顿可以忽略 p^3 和 $6p^2$ 而解 $10p-1=0$ 得到 $p=0.1$.由此得出 $y=2.1$ 是根的第二级近似.下一步是令 $p=0.1+q$ 并代入关于 p 的方程.在得到的方程 $q^3+6.3q^2+11.23q+0.061=0$ 中,两个最高次项又被忽略.解得到的线性方程得 $q=-0.0054$,得到 y 的一个新的近似是 2.0946.人们可以将该方法进行到任意远.牛顿本人又作了一步后停止,得到 $y=2.09455148$.他然后将数值方程的解法加以修正,使之适用于代数并计算了若干例.例如,$y^3+a^2y+axy-2a^3-x^3=0$ 的解由下式给出

$$y = a - \frac{x}{4} + \frac{x^2}{64a} + \frac{131x^3}{512a^2} + \frac{509x^4}{16384a^3} + \cdots.$$

而方程

$$y^5/5 - y^4/4 + y^3/3 - y^2/2 + y - z = 0$$

的解是

$$y = z + \frac{1}{2}z^2 + \frac{1}{6}z^3 + \frac{1}{24}z^4 + \frac{1}{120}z^5 + \cdots.$$

12.5.2 二项式定理

牛顿幂级数的发现得自他对沃利斯《无穷算术》的阅读,特别是确定圆面积的部分.实际上,他把沃利斯的著作发扬光大.在考虑面积时,沃利斯总是寻求特定的数值或者两个特定数值的比,因为他想确定两个固定值——如0和1——间的曲线下的面积.牛顿意识到如果人们计算从0到任意值 x 的面积就会看到进一步的规律,就是说,人们将曲线下的面积考虑为区间变动端点的一个函数.因此在看沃利斯计算圆面积这同一个问题时,他考虑了与沃利斯考虑的曲线相类似的一系列曲线,即曲线 $y=(1-x^2)^n$.但牛顿后来将这些曲线下的面积作为变量 x 的函数加以列表.例如,用现代的符号

$$\int_0^x (1-x^2)^0 \mathrm{d}x = x,$$

$$\int_0^x (1-x^2)^1 \mathrm{d}x = x - \frac{1}{3}x^3,$$

$$\int_0^x (1-x^2)^2 \mathrm{d}x = x - \frac{2}{3}x^3 + \frac{1}{5}x^5,$$

$$\int_0^x (1-x^2)^3 \mathrm{d}x = x - \frac{3}{3}x^3 + \frac{3}{5}x^5 - \frac{1}{7}x^7,$$

$$\int_0^x (1-x^2)^4 \mathrm{d}x = x - \frac{4}{3}x^3 + \frac{6}{5}x^5 - \frac{4}{7}x^7 + \frac{1}{9}x^9,$$

牛顿然后列出了 x 不同次幂的系数而不是数值面积:

$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	…	乘以
1	1	1	1	1	…	x
0	1	2	3	4	…	$-\frac{x^3}{3}$
0	0	1	3	6	…	$\frac{x^5}{5}$
0	0	0	1	4	…	$-\frac{x^7}{7}$
0	0	0	0	1	…	$\frac{x^9}{9}$

像沃利斯一样,牛顿意识到这里存在帕斯卡三角形,所以他试图作插值.实际上,要解圆面积的问题,他需要对应着 $n=1/2$ 的列的数值.为找到这些值,他重新发现了适用于正整数值 n 的帕斯卡公式

$$\binom{n}{k} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

并决定即使当 n 不是正整数时也使用这同一个公式.因此列中的项将是

$$\binom{\frac{1}{2}}{0} = 1,\quad \binom{\frac{1}{2}}{1} = \frac{1}{2},\quad \binom{\frac{1}{2}}{2} = \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)}{2} = -\frac{1}{8},$$

$$\binom{\frac{1}{2}}{3} = \frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\left(\frac{1}{2}-2\right)}{6} = \frac{1}{16},\cdots.$$

牛顿此时可以对任意的正数 k 填出对应于 $n=k/2$ 的表列.他进一步意识到在原来的表中每一项都

是它左边和上边的项之和.在他有了额外的插值列的表中,他略微修正该法则为每一项都是它左边第二行和它上边的项之和,用二项式系数公式找到的新项也满足这一法则.这不仅给了牛顿他插值正确的信心,而且还使他有信心在左边添加上了对应着负数 n 的列.求和法则使他清楚在 $n=-1$ 的列的第一个数一定是 1,而下一个一定是 -1,因为 $1+(-1)=0$,而 0 是 $n=0$ 一列的第二项.同样,$n=-1$ 的列的第三个数是 1,第四个是 -1,以此类推.当然,二项式系数公式也得出这些交替的 1 和 -1.因此牛顿计算从 0 到 x 的 $y=(1-x^2)^n$ 下的面积的插值表如下:

$n=-1$	$n=-\frac{1}{2}$	$n=0$	$n=\frac{1}{2}$	$n=1$	$n=\frac{3}{2}$	$n=2$	$n=\frac{5}{2}$	$\cdots$	乘以
1	1	1	1	1	1	1	1	$\cdots$	x
-1	$-\frac{1}{2}$	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2	$\frac{5}{2}$	$\cdots$	$-\frac{x^3}{3}$
1	$\frac{3}{8}$	0	$-\frac{1}{8}$	0	$\frac{3}{8}$	1	$\frac{15}{8}$	$\cdots$	$\frac{x^5}{5}$
-1	$-\frac{5}{16}$	0	$\frac{3}{48}$	0	$-\frac{1}{16}$	0	$\frac{5}{16}$	$\cdots$	$-\frac{x^7}{7}$
1	$\frac{35}{128}$	0	$-\frac{15}{384}$	0	$\frac{3}{128}$	0	$-\frac{5}{128}$	$\cdots$	$\frac{x^9}{9}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$

牛顿很快便意识到,首先,没有必要只处理分母为 2 的分数.对 $\binom{n}{k}$ 成立的乘法法则一定适用于 n 为任何分数的情形,不论是正还是负.第二,正如他在 1676 年 10 月 24 日给莱布尼茨的信中提到的,他意识到对 n 为整数时 $(1-x^2)^n$ 的项"可以以和它们产生的面积相同的方式进行插值;并且为这一目的只需要在表示面积的项中舍弃掉分母为 1,3,5,7 等等的项"[36](并且,当然,将相应的幂减 1).最后,没有理由将自己限制在形如 $1-x^2$ 的二项式.经过适当的修正,适用于任何 n 的幂级数 $(a+bx)^n$ 的系数都可以用二项式公式加以计算.因此,虽然基本上没有证明,牛顿发现了广义的二项式定理.但是,他完全相信它的正确性,因为在几种情况下它得到了他用其它方法得到的同样的答案.例如,牛顿注意到用除法得到的 $1/(1+x)$ 的级数与取指数为 -1 用二项式定理得到的级数相同:

$$(1+x)^{-1}=1+(-1)x+\frac{(-1)(-2)}{2!}x^2+\frac{(-1)(-2)(-3)}{3!}x^3+\cdots$$
$$=1-x+x^2-x^3+\cdots.$$

牛顿利用他关于 $y=1/(1+x)$ 下的面积是 $1+x$ 的对数的知识,通过对上面的级数逐项积分得出了 $\log(1+x)$ 的幂级数,进而将 $1\pm0.1, 1\pm0.2, 1\pm0.01$ 和 1 ± 0.02 的对数计算到小数后 50 位.利用适当的恒等式,例如 $2=(1.2\times1.2)/(0.8\times0.9)$ 和 $3=(1.2\times2)/0.8$ 以及对数的基本性质,牛顿能够计算许多小正整数的对数.

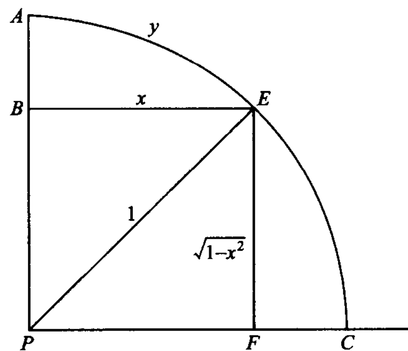


图 12.22　牛顿对 $y=\arcsin x$ 的幂级数.

二项式的知识使牛顿处理了许多其它有趣的级数.例如,他用几何推导得出了 $y=\arcsin x$ 的级数:假定圆 AEC 的半径是 1 而 $BE=x$ 是弧 $y=AE$ 的正弦,或者说 $y=\arcsin x$(图 12.22).已知扇形 APE 的面积是 $0.5y=$

0.5arcsin x.另一方面,它也等于从0到x的$y=\sqrt{1-x^2}$下的面积减去$0.5x\sqrt{1-x^2}$.根据他以前的计算,牛顿知道

$$\sqrt{1-x^2}=1-\frac{1}{2}x^2-\frac{1}{8}x^4-\frac{1}{16}x^6-\cdots,$$

通过逐项积分并给上面的级数乘以x,可以得到

$$y=\arcsin x=2\int_0^x\sqrt{1-x^2}\mathrm{d}x-x\sqrt{1-x^2}=x+\frac{1}{6}x^3+\frac{3}{40}x^5+\frac{5}{112}x^7+\cdots.$$

牛顿然后用他解隐性方程的方法解这个“方程”求得$x=\sin y$.由此欧洲数学在《分析学》中第一次出现了级数

$$x=\sin y=y-\frac{1}{6}y^3+\frac{1}{120}y^5-\frac{1}{5040}y^7+\cdots,$$

以及$x=\cos y$的级数,后一级数是牛顿通过计算$\sqrt{1-(\sin y)^2}$得到的.

今天人们在处理幂级数时,总会考虑收敛问题.似乎牛顿并不怎么担心这个问题.他在《分析学》临近末尾的地方写道:“无论对由有限项组成的方程执行了什么操作,…… 该方法(级数的)总可以对无限方程执行.当然,后者中的推理同其他的推理同样确定,它的方程同样精确,虽然我们人类只有有限的智力,既不能指明它们所有的项,也不能掌握它们到如此程度,以至于能从它们那里精确地得到我们要求的数量.”[37] 但是,牛顿清楚地意识到了他方法的限制,至少是直觉地意识到了这点.虽然他从未对收敛问题正式作出处理,他确实注意到,例如,在他得到双曲线$y=1/(1+x)$下的面积的计算过程中,这个对数数列的前面几项“是有相当用处的,并在x明显小于1时可以足够准确.”[38]

12.5.3 计算流数的方法

级数对牛顿的微积分十分重要.他利用它们处理所有不能表示为单变量的有限多项式的代数关系或者超越关系.但在他的《论方法》中当然包括更多的东西,《论方法》开始于他在给莱布尼茨第二封信中仅用密码作过暗示的问题,也是他认为是微积分两个基本方面的问题:“1.连续地给出距离的长度(就是说,在任何时间的),求任何指定时间的运动的速度.2.连续地给出运动的速度,求在任何指定时间走过的距离.”[39] 对牛顿说来,微积分的基本思想是同运动有关的.一个方程中的所有变量都被看作是 —— 至少是隐含地 —— 依赖于时间的距离.当然,这一思想不是牛顿创造的,但他使得运动的思想成为主导思想:“我把量看成好像是当运动的物体绘出轨迹时由连续增加的距离产生的.”[40] 牛顿实际上把时间的稳定增加本身看作是一条公理,因为他没有定义时间.他作过定义的是**流数**的概念:依赖于时间的量x(称为**流变**)的流数$\dot{x}$是x在生成运动中增加的速度.在这一早期著作中,牛顿没有试图给对速度作进一步的定义.牛顿相信连续变化着的运动的概念是完全直觉的.

牛顿用一种完全直截了当的算法解决了问题1,该算法确定通过形如$f(x,y)=0$的方程联系起来的两个流变x和y的流数$\dot{x}$和$\dot{y}$的关系:“按照某一流变,比如x,的次数将表达给定关系的方程加以排列,用它的项乘以一算术数列再乘以$\frac{\dot{x}}{x}$.对每个流变分别施加这一运算再令所有这些积的和为0,你便得到需要的方程.”[41] 牛顿提出方程$x^3-ax^2+axy-y^3=0$作示例.首先把它考虑成一个关于x的3次多项式.牛顿用数列3,2,1,0去乘,得到$3x^2\dot{x}-2ax\dot{x}+ay\dot{x}$.然后把方程考虑成一个关于$y$的3次多项式并使用同样的数列,他计算出$ax\dot{y}-3y^2\dot{y}$.令和等于0得到需要的关系$3x^2\dot{x}-$

$2ax\dot{x}+ay\dot{x}+ax\dot{y}-3y^2\dot{y}=0$,用比来表示,这个结果是 $\dot{x}:\dot{y}=(3y^2-ax):(3x^2-2ax+ay)$.

在牛顿计算流数的法则中有几个重要思想要注意.首先,牛顿不是计算导数,因为他一般不从函数开始.他实际计算的是由给定方程确定的曲线满足的微分方程.换言之,给定 $f(x,y)=0$,式中 x 和 y 都是 t 的函数,牛顿的程序产生的结果今天可以写为:

$$\frac{\partial f}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\partial f}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}=0.$$

第二,牛顿采用了胡德用任意算术数列相乘的概念.但在实践中,牛顿一般采用以流变最高幂开始的数列.第三,如果 x 和 y 都被看作是 t 的函数,那么现代的导数乘积法则就被嵌入了牛顿的算法.任何包含 x 和 y 的项都被乘了两次,两项相加.

牛顿通过无穷小量说明了他法则的合理性.他首先定义一个流变的瞬(moment)是在"无穷小"的时间里它增加的量.例如,x 在无穷小时间 o 的增加是 x 的速度和 o 的乘积,或者 $\dot{x}o$.由此得出经过这一时间段,x 将变为 $x+\dot{x}o$,类似地 y 变为 $y+\dot{y}o$."结果,一个表示在所有时间都不变化的流变的关系的方程将同样表示 $x+\dot{x}o$ 和 $y+\dot{y}o$ 间的关系以及 x 和 y 的关系;因此 $x+\dot{x}o$ 和 $y+\dot{y}o$ 就可以在所说的方程中代换后面的量 x 和 y."[42]

牛顿接着说明该方法如何应用到早先给出了例子 $x^3-ax^2+axy-y^3=0$.用 $x+\dot{x}o$ 代换 x 并用 $y+\dot{y}o$ 代换 y,新的方程变为

$$\begin{aligned}&(x^3+3x^2\dot{x}o+3x\dot{x}^2o^2+\dot{x}^3o^3)-(ax^2+2ax\dot{x}o+a\dot{x}^2o^2)\\&\quad+(axy+ay\dot{x}o+ax\dot{y}o+a\dot{x}\dot{y}o^2)-(y^3+3y^2\dot{y}o+3y\dot{y}^2o^2+\dot{y}^3o^3)=0.\end{aligned}$$

"现在根据假设 $x^3-ax^2+axy-y^3=0$,在这些项被删除并把余下的项除以 o,则还剩下

$$3x^2\dot{x}+3x\dot{x}^2o+\dot{x}^3o^2-2ax\dot{x}-a\dot{x}^2o+ay\dot{x}+ax\dot{y}+a\dot{x}\dot{y}o-3y^2\dot{y}-3y\dot{y}^2o-\dot{y}^3o^2=0.$$

但进一步,因为 o 被假定为无穷小,所以它才能够表示量的瞬,包含它作为因子的项相对于其它项将等价于0.因此我把它扔掉,还剩下 $3x^2\dot{x}-2ax\dot{x}+ay\dot{x}+ax\dot{y}-3y^2\dot{y}=0$,正如……上面."[43]

尽管这一计算仅仅是个例子而不是证明,牛顿注意到它可以立即被推广:"因此可以观察到没有乘 o 总是会消失,那些被 o 一次以上的幂相乘的项也一样;而剩下的项在除过 o 之后将总是具有它们根据法则应该具有的形式.这就是我想说明的."[44] 换言之,牛顿假定读者意识到 $(x+\dot{x}o)^n$ 的展开式中 $x^{n-1}\dot{x}o$ 的系数是 n 本身.但也请注意牛顿"扔掉"任何中间有 o 的项这一步的惟一理由是它们"相对于其它项将等价于0".这里没有极限论证.这里只有对无限小时间增量的性质的直觉概念.

导数的乘积法则实际上被嵌入了牛顿的算法.牛顿对现代链式法则的处理是通过代换进行的.例如,为确定方程 $y-\sqrt{a^2-y^2}$ 中流数的关系,他用 z 代替平方根并处理方程 $y-z=0$ 和 $z^2-a^2+x^2=0$.第一个给出 $\dot{y}-\dot{z}=0$ 而第二个给出 $2z\dot{z}+2x\dot{x}=0$ 或者 $\dot{z}=-x\dot{x}/z$.因此 x 和 y 的流数间的关系为

$$\dot{y}+\frac{x\dot{x}}{\sqrt{a^2-x^2}}=0.$$

同样的方法对处理商也成立.

为解决给定流数间的关系求流变间关系的问题,牛顿在可能的时候简单地把上述过程反过来:"因为该问题是前一问题的逆问题,它应该用相反的方式来解决:即根据 x 的次数排列与 $\dot{x}$ 相乘的项,并除以 $\dot{x}/x$ 再除以次数,……对与 $\dot{y}$ 相乘的项执行同样的运算,抛弃掉冗余的项,令所得项的和为0."[45] 他选取了前面用到的同一个问题作为例子.从 $3x^2\dot{x}-2ax\dot{x}+ay\dot{x}-3y^2\dot{y}+ax\dot{y}=0$ 开始,他用 $\dot{x}/x$ 除以含有 $\dot{x}$ 的项(或者与此殊途同归,删去 $\dot{x}$ 并把 x 的幂增加1),然后再把每项除以 x 的

新的次数得到 $x^3 - ax^2 + axy$.对包含 $\dot{y}$ 的项施以类似的运算,他得到 $-y^3 + axy$.注意到 axy 出现了两次,他删除掉一个并得到最后的方程 $x^3 - ax^2 + axy - y^3 = 0$.

牛顿当然意识到这一过程并非总能行得通.他实际上建议人们始终检查结果.但假如问题不能用这一简单的"反导数"解决,牛顿通常使用幂级数的方法.因为流数方程 $\dot{y} = x^n\dot{x}$ 或者 $\dot{y}/\dot{x} = x^n$ 确定的流变方程是 $y = x^{n+1}/(n+1)$,他建议当 $\dot{y}/\dot{x}$ 仅依赖于 x 时,人们应该把比值表示为幂级数并对每一项应用法则.例如,方程 $\dot{y}^2 = \dot{x}\dot{y} + x^2\dot{x}^2$ 可以改写为 $\dot{y}^2/\dot{x}^2 = \dot{y}/\dot{x} + x^2$.这个关于 $\dot{y}/\dot{x}$ 的二次方程可以被求解得出 $\dot{y}/\dot{x} = 1/2 \pm \sqrt{1/4 + x^2}$.应用二项式定理,人们得到两个级数:

$$\frac{\dot{y}}{\dot{x}} = 1 + x^2 - x^4 + 2x^6 - 5x^8 + \cdots \quad \text{或} \quad \frac{\dot{y}}{\dot{x}} = -x^2 + x^4 - 2x^6 + 5x^8 + \cdots.$$

原问题的解就能轻易地被找到,是

$$y = x + \frac{1}{3}x^3 - \frac{1}{5}x^5 + \frac{2}{7}x^7 + \cdots \quad \text{或} \quad y = -\frac{1}{3}x^3 + \frac{1}{5}x^5 - \frac{2}{7}x^7 + \cdots.$$

如果 $\dot{y}/\dot{x}$ 由包括 x 和 y 的方程给出,解法将更为复杂,但即使在那时,牛顿的基本思想还是将给定的方程用幂级数表示.

12.5.4 流数的应用

完成了流数的计算,牛顿用它们解决各种各样的问题.通过令相关的流数为 0 可以求得最大值和最小值.因为"当一个量最大或最小时,它的流在那一时刻既不增加也不减小;因为如果它增加,这证明它较小并且将立即变得比现在大;而相反的话,则它会减小."[46] 他再次使用了方程 $x^3 - ax^2 + axy - y^3 = 0$ 作为例子来确定 x 的最大值.在含有流数的方程中令 $\dot{x} = 0$ 他得到 $-3y^2\dot{y} + ax\dot{y} = 0$ 或者 $3y^2 = ax$.然后同时解这个方程与原方程来求得所要求的 x 的值.同样,要求 y 的最大值,人们令 $\dot{y} = 0$ 并使用得到的方程 $3x^2 - 2ax + ay = 0$.但牛顿对该方法的讨论是简短的,并且他没有给出判断求得的值是最大值还是最小值的标准.据估计,这一判断可以根据给定问题的上下文作出.

牛顿作切线的中心思想是利用巴罗的微分三角形.因此,如果 x 变为 $x + \dot{x}o$ 而 y 变为 $y + \dot{y}o$,则这一三角形的边的比值 $\dot{y}o : \dot{x}o = \dot{y} : \dot{x}$ 就是切线的斜率,该斜率被视为是描画曲线的粒子瞬时运动的方向.这一比值又等于纵坐标 y 与次切距 t 的比值.因为作切线意味着找次切距,牛顿简单地注意到 $t = y(\dot{x}/\dot{y})$.作为在这一计算和其它计算的一个略微简化,牛顿有时候令 $\dot{x} = 1$.这与把 x 考虑为均匀流动相等价,或者说它自己就表示时间.

牛顿对流数运用的最后一个例子是他对曲线曲率的计算,这是一个"带有异常美妙并对曲线科学极其有用的印记"[47] 的问题.牛顿用圆来定义曲率,即,他注意到一个圆在各处都有同样的曲率并且两个圆的曲率同它们的半径成反比.用现代的术语说,半径为 r 的圆的曲率被定义为 $\kappa = 1/r$.对一任意曲线,牛顿将一点处的曲率定义为在该点与曲线相切并且在曲线和圆之间作不出其它相切圆的圆的曲率.该定义意味着点 D 处的这一**密切圆**还通过与 D 无限接近的点 d.为求点 D 处的曲率,人们只需要找到该圆的半径,就是说,到法线与曲线分别在 D 和 d 处交点的距离 DC(图 12.23).过 D 和 d 作曲线的切线 dDT,完成矩形 $DGCH$,在 GC 上取点 g 使得 $Cg = 1$,令 $AB = x$,$BD = y$ 且 $g\delta = z$,牛顿从三角形 DBT 与三角形 $Cg\delta$ 的相似得出 $cg : g\delta = TB : BD$ 或者 $1 : z = \dot{x} : \dot{y}$.因为 d 和 D 无限接近,可得 $\delta f = \dot{z}o$,$DE = \dot{x}o$ 且 $dE = \dot{y}o$.并且因为人们可以把 DdF 看作是直角三角形,$DE : dE = dE : EF$ 且因此 $DF = DE + EF = DE + (dE^2/DE) = \dot{x}o + (\dot{y}^2o/\dot{x})$.所以,$Cg : CG = \delta f : DF = \dot{z}o : (\dot{x}o + \dot{y}^2o/\dot{x})$ 且 $CG = (\dot{x} + \dot{y}^2)/\dot{x}\dot{z}$.假定 $\dot{x} = 1$,牛顿得到 $\dot{y} = z$,

$$CG = (1 + z^2)/\dot{z}, \quad DG = (CG \cdot BD)/BT = Z \cdot CG = (z + z^3)/\dot{z},$$

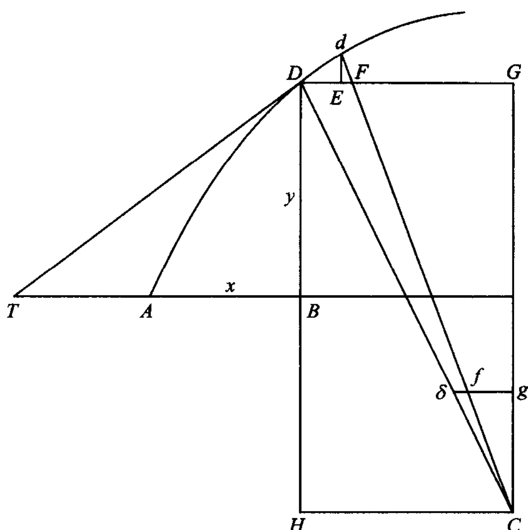


图 12.23　牛顿求曲率的方法.

最后

$$DC = \sqrt{CG^2 + DG^2} = \frac{(1 + z^2)^{3/2}}{\dot{z}}.$$

牛顿的结果自然等价于现代的形式,即 $y = f(x)$ 的曲率为

$$\frac{y''}{(1 + y'^2)^{3/2}},$$

因为他的 z 是我们的 y',他的 $\dot{z}$ 是我们的 y''.

12.5.5　求面积的步骤

《论方法》的第二个问题是给定速度求距离.牛顿在他的研究中很早就意识到该问题等价于根据曲线的方程求曲线下的面积.从他所读的沃利斯的书中,他也知道如何求方程为有限多个形如 $ax^n(n \neq -1)$ 的项的和的曲线下的面积并把这一思想推广到无穷和或者说级数上.但是,牛顿还发现并且应用了微积分基本定理来求解面积问题.对他来说,该定理实际上是自明的.因为他把曲线 AFD 看作是由 x 和 y 的运动产生的,由此推知面积 $AFDB$ 是由动坐标 BD 产生的(图 12.24).因此,显而易见,面积的流数实际上是纵坐标与 BD 流数的乘积.就是说,如果 z 代表曲线下的面积,则 $\dot{z} = y\dot{x}$ 或 $\dot{z}/\dot{x} = y$.该方程立即可以解说为现代基本定理的一部分,如果 $A(x)$ 代表从 0 到 x 的 $y = f(x)$ 下的面积,则 $\mathrm{d}A/\mathrm{d}x = f(x)$.牛顿注意到面积 z 可以使用已经讨论过的求流量的技巧明确地从方程 $\dot{z}/\dot{x} = y$ 求出.但是正如他在几页后所写的"至今我们已经阐明了通过将其化为包含无穷多简单项的方程对由比较复杂的方程定义的曲线求积.但是,这类曲线有时候也可以用有限方程的手段求面积."[48]

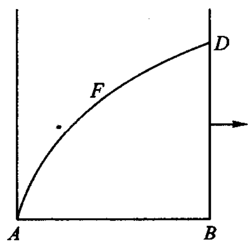


图 12.24　牛顿和基本定理.

为了用有限方程求曲线面积(即求曲线下的面积),人们需要一张积分表.牛顿提供了一份相当丰富的积分表.他的表中的第一项是那个简单的 —— 曲线 $y = ax^{n-1}$ 下的面积是 $\frac{a}{n}x^n$ —— 但其余的则复杂得多.在这个牛顿的表的一小段摘录中,右边的函数 z 表示左边的函数 y 下的面积:

$$y = \frac{ax^{n-1}}{(b + cx^n)^2}, \qquad z = \frac{(a/nb)x^n}{b + cx^n},$$

$$y = ax^{n-1}\sqrt{b + cx^n}, \qquad z = \frac{2a}{3nc}(b + cx^n)^{3/2},$$

$$y = ax^{2n-1}\sqrt{b + cx^n}, \qquad z = \frac{2a}{nc}\left(-\frac{2}{15}\frac{b}{c} + \frac{1}{5}x^n\right)(b + cx^n)^{3/2},$$

$$y = \frac{ax^{2n-1}}{\sqrt{b + cx^n}}, \qquad z = \frac{2a}{nc}\left(-\frac{2}{3}\frac{b}{c} + \frac{1}{3}x^n\right)\sqrt{b + cx^n}.$$

同现代的积分表相比,人们注意到牛顿的表没有列出超越函数,没有正弦、余弦甚至没有对数.尽管牛顿知道这些函数的幂级数,他从未将它们与代数函数同等看待.他没有通过将正弦、余弦和对数同多项式和其它的代数表达式结合起来的方式处理过它们.但是,牛顿确实将他的表扩展到了那些其积分在今天要用超越函数来表示的函数,他将这些积分表示为由某些圆锥曲线围成的面积,这些面积可以用幂级数的技巧来计算.例如,设 $y = x^{n-1}/(a + bx^n)$,他写道面积 z 可以表示为双曲线 $v = 1/(a + bu)$ 下面积的 $1/n$,此处 $u = x^n$.

但是他的表对一些曲线是无法给出答案的,即那些用几何方法定义的曲线,例如旋轮线.给定当圆 ALE 沿 EF 滚动时点 A 绘出的旋轮线 ADF(图 12.25),牛顿注意到旋轮线在任意一点 D 处的切线 DT 始终平行于 AL,此处 L 是过 D 平行于 EF 的直线与圆的交点.这是因为产生旋轮线的运动是由直径 AE 沿 EF 运动和 A 绕圆周相等运动组成的.因为 $\dot{y}:\dot{x}$ 是切线的斜率,而这等于 $DG:GT$ 或者 $DG:BL$,由此得出 $(DG)\dot{x} = (BL)\dot{y}$.但 $(DG)\dot{x}$ 是面积 ADG 的流数而 $(BL)\dot{y}$ 是面积 ALB 的流数.由此得出旋轮线半拱下的面积 AHF 等于半圆 $ALEB$ 的面积.因为矩形 $AEFH$ 是生成圆面积的两倍,由此得出旋轮线一个完整拱上的面积是生成圆面积的 3 倍,罗伯华曾在 30 多年前发现了同样的结果.

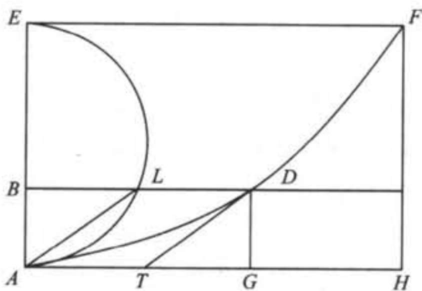


图 12.25 牛顿对旋轮线一拱所围成的面积的计算.



图 12.26 为纪念牛顿的《原理》发表三百周年而发行的英国邮票.

牛顿的《论方法》中还有许多其它内容,包括相当于现代的代换法则(如在前面的积分例子中的)的技巧、分部积分法以及求弧长的方法.因此,这一从未发表过的著作实际上包含了任何现代微积分教程最初几章里所有重要的思想,还有一些在那里被认为是过于高深的思想.但缺少的一个思想是极限的思想.这并不是说牛顿从未考虑过这一思想.他考虑过,但仅在同他的代表作 1687 年的《自然哲学的数学原理》有关场合才发表这些思想,他在《自然哲学的数学原理》中表述了他的运动定律并用这些运动定律、引力理论还有宇宙的数学得到了他的"世界体系"(图 12.26).

12.5.6 《论运动》和宇宙的数学

宇宙的数学的故事开始于 1684 年的夏天，当时，年轻的英国天文学家爱德蒙·哈雷(Edmond Halley, 1654—1741)旅行到剑桥并向牛顿提出了一个关键的问题："假定朝太阳的引力与行星到太阳的距离的平方成反比，行星走出的曲线会是什么样？"[49] 牛顿即刻的回答是他已经"计算"过这一问题的答案，曲线将会是椭圆. 当哈雷向牛顿逼问细节时，这位剑桥教授许诺不久后寄给他. 好几个月过去了，直到 1684 年 11 月哈雷才收到牛顿寄来的一份 10 页的论文，该论文的目的不仅是回答原来的问题，还勾画出了用力的术语改写天文学的轮廓. 哈雷被《论天体在轨道里的运动》这本著作深深打动，于是他急忙回到剑桥试图劝说牛顿出版该著作. 显然，哈雷并不需要花费很大力气来说服牛顿，牛顿已经着手将该论文修订和扩展成《原理》.

即使是今天，《原理》仍是一本使人感到畏惧的书，部分原因是因为它是用几何的语言写成的，而不是我们现代的分析语言. 因此，我们将考虑《论运动》这一对力、加速度和天体运动更为基本的介绍，而不去分析《原理》中的数学. 物理和数学的基本思想在《论运动》里都有，只是缺少代表作中大量的细节. 这个短小的论文包含有 4 个定理和 7 个问题，并且像平常一样是从若干定义和"假设"开始的. 定义涉及两种类型的力，在当时即使牛顿对力的概念也不清楚. **向心力**是"吸引"物体到某一个被视为中心的点的力，而物体**自身固有的力**是使物体努力维持直线运动的力. 注意这里甚至惯性都需要一个力. 另一方面，牛顿的一个假设是"任何仅受惯性力的物体继续沿直线匀速运动直到无穷，除非它被某个外在的东西所阻挡."[50] 牛顿再利用另一个假设"给定时间内一物体在一组合力的作用下与这些分力在同样的时间次第作用下到达的位置相同"就可以处理在轨道里运动的天体了. 他的基本思想是向心力和沿切线的天体惯性力的联合作用产生了运动的实际曲线. 最后，为提供一种量度力的效果的方法，牛顿提出"在任何向心力作用下物体在运动的最初走过的距离同时间的平方成比例". 这一假设采用了伽利略关于靠近地球表面的落体走过的距离与时间平方成正比的定理并将其推广到变力. 但推广仅在"无穷小的情况下"成立，或者如牛顿所说，仅在运动的"最开始"成立.

我们考虑牛顿的头一个定理是"所有在轨道里运行的天体的通过连接中心的半径所扫过的面积同时间成正比." 该结果当然是开普勒第二定律. 证明显示了牛顿对其中包含的物理原理以及无穷小数学的理解. 牛顿从把时间分成有限个相等的部分开始. 假定在第一个时间段里，天体在"惯性力"作用下沿直线段 AB 运动(图 12.27). 如果没有什么妨碍它，在下一个时间段它会沿同一方向上一个相等的线段 Bc 运动. 因此如果从 A、B 和 c 向中心 S 引线段，三角形 ASB 和 BSc 将有相等的面

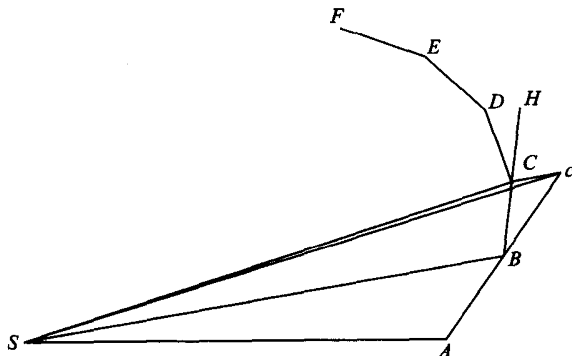


图 12.27　牛顿对面积的确定.

积.但是因为天体被拉向中心,牛顿假定当它到达 B 时,向心力发生作用并使得天体改变路线并沿 BH 方向运动.现在从 c 引平行于 BS 的直线交 BH 于 C.根据牛顿关于合力的假设,在第二个时间段的结束,天体将出现在 C.如果我们现在连接中心 S 和 C,则三角形 BSC 同三角形 BSc 面积相等因此同三角形 ASB 的面积也相等.因为该论证可以在其它相等的时间段重复,由此可得在同样的时间内走过同样的面积,至少在向心力间断起作用的假设下如此.牛顿当然知道力是持续起作用的.因此,他这样结束他的证明:"令这些三角形有无数个并且无穷小,以至于每个三角形对应着一个时刻,再有向心力的持续作用,命题将成立."[51]"向极限的过渡",像我们今天会说的,在牛顿这里很轻易就完成了.

为了处理向心力和轨道的关系,牛顿需要一个量度力的几何方法.这点他是在定理 3 中完成的,该定理中天体是以任意曲线在以 S 为中心的轨道中运行的(图 12.28).如果 PX 同曲线在点 P 相切,QT 与 PS 垂直,其中 Q 为轨道上另一点 Q,如果向 PX 引 QR 平行于 PS,则"当人们总是取立体在 P 和 Q 重合时的最终量"[52],向心力将同立体元$(SP^2 \times QT^2)/QR$ 成反比.牛顿的证明假定同切线的偏离 QR 与力成正比.(这一假设后来成为牛顿运动第二定律,运动的改变同力成正比.)从他关于伽利略结果的假设,他还知道 QR 同时间的平方成正比,或者根据第一定理,同扫过的面积的平方成正比.因此 QR 同向心力和$(SP \times QT)^2$ 都成正比,结果因此成立.但再次注意该结果仅在"无穷小的情况下" 成立,或者用现代的术语,"在极限情况" 成立,牛顿关于这些思想的直觉使得他在没有使用任何极限理论的工具的情况下就得到了结果.

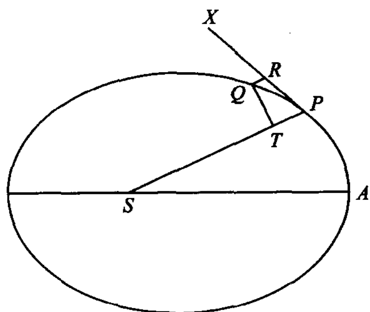


图 12.28 确定度量力的几何方法.

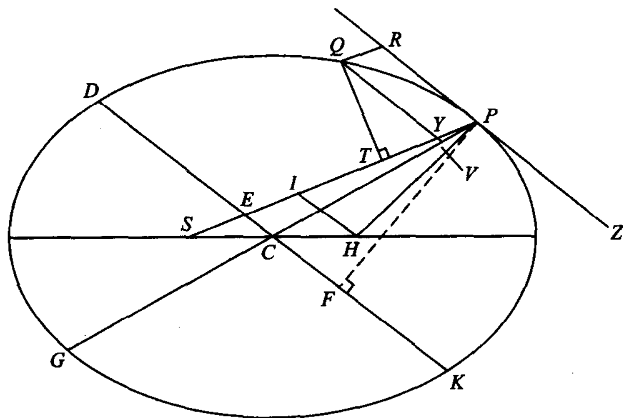


图 12.29 椭圆轨道服从平方反比的力的定律.

手头有了对力的几何表示,牛顿现在可以对各种特定的轨道计算力了.他对圆周轨道计算了力指向圆周上一点的情况,对椭圆轨道计算了力指向中心的情况,最后 —— 最有意思的一种情况 —— 对椭圆轨道计算了力指向一个焦点 S 的情况(图 12.29).在后一种情况中,采用与定理 3 中同样的符号,令 DK、PG 是椭圆的共轭直径,图中 DK 平行于 PR,作 QV 平行于 PR 交 PG 于 V.并且,牛顿作 SP 分别交 DK、QV 于 E、Y,这样完成了平行四边形 $QYPR$.像往常一样,我们用 a 和 b 分别表示椭圆长半轴和短半轴的长,p 表示参量.牛顿接着证明了一个引理,是 $PE = a$.因为如果 H 是椭圆的第二个焦点,同时作 HI 平行于 EC,则 $ES = EI$ 且 $\angle PIH = \angle YPR = \angle ZPH = \angle PHI$.因此 $PE = (PS + PI)/2 = (PS + PH)/2 = 2a/2 = a$.他然后给出了 5 个比例式,最后他将把它们相乘得到他最终的结果:

$$p \times QR : p \times PV = QR : PV = PY : PV = PE : PC = a : PC, \tag{12.1}$$

$$p \times PV : GV \times PV = p : GV, \tag{12.2}$$

$$GV \times PV : QV^2 = PC^2 : CD^2, \tag{12.3}$$

$$QV^2 : QY^2 = M : N, \tag{12.4}$$

$$QY^2 : QT^2 = PE^2 : PF^2 = a^2 : PF^2 = CD^2 : b^2. \tag{12.5}$$

根据三角形 PVY 与 PCE 相似以及引理,比例 12.1 成立,而比例 12.2 不过是消去律.牛顿从他对圆锥曲线的研究知道了比例 12.3;这是阿波罗尼乌斯关于一对共轭轴的命题 Ⅰ - 21.(见第三章,习题 25d.)命题 12.4 不过是 M 和 N 的定义,而最后一个比例取决于三角形 QTY 与 PFE 相似以及阿波罗尼乌斯的 Ⅶ - 31,即任何一对共轭直径上所作的矩形面积相等.(见见第三章,习题 25e.)如果人们现在将这些比例乘在一起,记得 $b^2 = pa/2$,结果是 $p \times QR : QT^2 = (2PC : GV) \times (M : N)$. 但是当 P 和 Q"融合"时,右边的两个比值成为等量的比.由此得出 $p \times QR = QT^2$ 并且两边同乘以 SP^2/QR,得

$$p \times SP^2 = \frac{SP^2 \times QT^2}{QR}.$$

因此,向心力同 $p \times SP^2$ 成反比,并且因为 p 是常量,同点到焦点的距离 SP 的平方成反比.

虽然该结果处理的是行星绕太阳的实际轨道,但牛顿没有回答哈雷的问题.他回答了相反的问题,即如果轨道是椭圆,则力同距离的平方成反比.但非常奇怪,在对该命题的一段评论中,牛顿确认了开普勒第一定律:"所以,主要的行星沿以太阳的中心为一个焦点的椭圆轨道运行."[53] 他混淆了定理和逆定理的区别吗?在过去的三个世纪里对这个问题的答案连篇累牍.最初是约翰 · 伯努利,他认为牛顿搞混了,并作出了平方反比的力将产生椭圆轨道的第一个解析证明.但有趣的是,牛顿在《论运动》和《原理》中提供了使得伯努利能给出证明的工具.在《论运动》中,如果天体在一个与距离平方成反比的力的作用下并且运行的轨道确实是圆锥曲线时,牛顿说明了如何确定一个天体运行的特定轨道.在《原理》中,他给出了若干说明在任何向心力作用下计算轨道的通用方法的命题.

从现代的观点看,证明平方反比的力产生椭圆轨道需要求解一个微分方程,而牛顿在《原理》中的结果正是给出了如何作到这点的轮廓,如果,如他所说,"曲线图形的求积"即对某些函数的积分.但到《原理》第 2 版特别是第 3 版出版时(1713 和 1726),牛顿意识到必须说更多的话.因此他简短地证明了以下结果:可以通过一个给定点作焦点、切线和曲率为给定值的圆锥曲线,圆锥曲线上的运动满足力的平方反比律,并且因为力和速度共同决定曲率,该圆锥曲线是平方反比律要求的初值问题的惟一解.[54]

但牛顿的确在《论运动》中证明了开普勒第三定律,使用的依然是无穷小方法.因此牛顿在这篇短文中勾勒了《原理》的重要结果.他从对向心力的一个假设出发证明(基本上)了开普勒的三个定律并且说明了椭圆轨道要求一种特殊类型的力,该力在《原理》的第三卷中成为了牛顿万有引力定律的基础.他没有完全做成功的一点是在他的数学过程中建立起我们今天要求的严格性.

12.5.7 《原理》和极限的概念

牛顿在《论运动》中使用了类似于极限的论证,此前他还在创立流数时使用过此类论证.但是,只是在《原理》中,牛顿试图将这一概念精确化,他从头一卷的第一节开始:

引理 1　在任何连续的时间内连续地向等量收敛的一些数量以及数量的比,如果在那段时间终止前比任何给定的差值都更接近,最终将变得相等.

证明是显然的.如果这些量最终不相等,它们将相差一个正值 D,因此趋于相等的程度不能超

过 D,这是矛盾.他进而在他"消逝的量的最终比"的概念中应用了极限的思想.例如,在引理7中他证明了当两个端点彼此靠近时(或者当每个的长"消逝"时)一段弧和它上面的弦的最终比是等量的比.实际上,他在《原理》写作之前20年关于流数的工作中就在瞬时速度的构思中使用了这一概念,因为瞬时速度就是当距离和时间消逝时两个量的比.但牛顿也知道那些埋头于希腊几何的人会反对这一概念.在对第一节的评注中,他试图答复他的批评者:

> 也许人们会反对说,消逝的量不存在最终比;因为在这些量消逝之前比例不是最终的,而当它们消逝了,就没有了.但用同样的论证可以断定,一个到达某一位置并在那里停止的物体没有最终速度;因为在物体到达该位置之前速度不是最终的;当它到达之后,不存在速度.但答案是简单的;因为最终速度是指物体借以移动的速度,既不是在它到达最终位置运动停止之前也不是在以后,而是在到达的瞬间 …… 以同样的方式,消逝的量的最终比要被理解成既不是这些量消逝前也不是消逝后的比,而是它们借以消逝的比 …… 存在一个在运动结束时速度可以达到但不能超过的极限.这就是最终速度.对所有开始产生或结束的量和比也有类似的极限 …… 这些量借以消逝的最终比不是真正的最终量(即,没有不可分量时)的比,而是这些无穷减少的量的比始终收敛趋向的极限,它们比任何给定的差都更靠近这一极限,但永远不会超过,并且实际上在这些量无穷变小前也不能达到.[55]

牛顿的话翻译成代数命题将给出极限的一个定义,这个定义接近但不等价于现代的定义.牛顿从未作出这一翻译.但是,似乎很清楚,牛顿直觉地知道他在使用"极限"计算流数时做的是什么.他知道他可以处理包含他的消逝的 o 的量的比并且在最后令 o 等于0.他的答案是正确的,并且在《论运动》和《原理》中他能够把这些思想应用到物理学的基本原理上.

经常有人断言牛顿是为了在《原理》中提出世界的体系才创立了微积分.从他大量手稿中得到的证据表明这不是事实,实际上微积分是在物理之前很久就创立了.无论如何,正如从前面几节中可以清楚看到的,牛顿的确使用了微积分的思想和方法推导了许多物理结果,虽然使用的不是代数方法.在他以几何为基础的物理论证中,他通常分3步进行.首先,他就有限的区域建立起一个结果.接着,他假定结果在同种类的无限小区域仍然成立.最后,他利用无限小的结果得出关于原图形的一些结论.例如,在《论运动》的定理1中,先对有限的三角形证明了面积法则,然后他简单地假定这些结果对无限小三角形仍然成立.接着他似乎断言,因为面积法则对无穷小的情况成立而由轨道围成的区域由无穷多个无限小三角形组成,面积法则对整个区域必然成立.同样地,在定理3中,他说明力可以用一个特定的立体来度量,然后说明这一结果对无穷小的情况成立,最后在接下来的应用中,说明了无穷小的结果能够转化为处理有限距离的结果.并且在他流数的论证中,不论是代数的还是几何的,他都使用了同样的三个步骤.他首先使用有限量发现一个结果,然后断定即便某些量是无穷小结果仍然成立,最后将这个新结果应用到有限的情形.

补遗 12.3	牛顿、莱布尼茨和微积分的发明
	牛顿和莱布尼茨被认为是微积分的发明人,而不是费马、巴罗或其他人,原因是他们完成了4项任务.他们都提出了同微积分两个基本问题相联系的基本概念 —— 对牛顿是流数和流量,对莱布尼茨是微分和积分 —— 两个问题是极值和面积.他们都发展了使人们能方便地使用这些概念的符号和算法.他们理解并运用了他们两个基本概念的逆关系.最后,他们使用这两个概念解决了许多从前不能解决的困难的问题.但两个人都没有做的,是用经典希腊几何的严格性为他们的方法奠定基础,因为两人实际上都使用了无穷小量.

所以虽然牛顿不是为从事天体力学发明了微积分,但认清一点很重要,就是他采用流数的思想作为他主要的物理工作的数学基础.并且尽管他直到晚年才发表流数方面的任何论文,但是,当他在 17 世纪 80 年代中期开始构建他的世界体系时,他 17 世纪 60 年代中期在剑桥的工作室以及沃尔索普的家中产生的思想被证明是关键的.注意到下面一点也很重要,就是与牛顿本人在他晚年的回忆录中试图使我们相信的相反,《论运动》和《原理》的基本思想在 17 世纪 60 年代并没有被提出.他当然在那时已经开始思考引力的问题,但只有到了 17 世纪 80 年代他才能把数学和物理的思想融合到他 1687 年的巨作中.

《原理》这一有理由被认为是科学革命中最重要的著作在以后两百年中界定了物理学的学科范畴.它确保了牛顿的声誉并最终在 1696 年使他成为造币厂厂长,1703 年成为皇家学会会长.另一方面,牛顿的微积分的影响相对较小,因为只有一部分微积分被发表而且发表的这一部分也是在写成许多年后才面世的.实际上,在牛顿自己发现之后大约 8 – 10 年完成的工作才构成了第一篇微积分出版物的基础,这就是微积分的共同发明人戈特弗里德·威廉·莱布尼茨(Gottfried Wilhelm Leibniz,1646 – 1716)的工作(补遗 12.3).

12.6　戈特弗里德·威廉·莱布尼茨

正如在本章开头所指出的,惠更斯在他 1672 年至 1676 年在巴黎逗留期间鼓励莱布尼茨阅读诸如范·舒滕版的笛卡儿的《几何学》以及包含了微分三角形的帕斯卡的著作这类材料从而把他引上了数学研究的前沿.在这段时期的最后,莱布尼茨才能够开始将他引上独立发明微分和积分的研究.但大约只过了 10 年,他已经开始以短小注记的形式在他协助创立的德国科学杂志《教师学报》上发表他的结果了.这里对莱布尼茨微积分的叙说将取自于这些注记.莱布尼茨在 1714 年为回应英国数学家断言他窃取了牛顿的方法所写的一篇题为《微积分的历史和起源》的著作以及他在巴黎早期工作的手稿,这些手稿实际上是他记录关于新微积分思想的日志.[56]

12.6.1　和与差

使莱布尼茨创立微积分的思想是数的序列的和与差的逆关系.他注意到如果 A、B、C、D、E 是一个递增数列,而 L、M、N、P 是差的序列,则 $E - A = L + M + N + P$,就是说,"相邻项的差的和总等于序列中最前项与最末项的差,不论它们的数目有多大."[57] 由此可得,差的序列很容易求和.因此,莱布尼茨不仅考虑了帕斯卡的算术三角形,其中每一列包含了前一列元素的和,或者反过来,每一列包含了后一列的差,而且还考虑了有相似性质的一个分数的新三角形,他称其为"调和三角形":

$$
\begin{array}{lllllll}
\frac{1}{1} \\
\frac{1}{2} & \frac{1}{2} \\
\frac{1}{3} & \frac{1}{6} & \frac{1}{3} \\
\frac{1}{4} & \frac{1}{12} & \frac{1}{12} & \frac{1}{4} \\
\frac{1}{5} & \frac{1}{20} & \frac{1}{30} & \frac{1}{20} & \frac{1}{5} \\
\frac{1}{6} & \frac{1}{30} & \frac{1}{60} & \frac{1}{60} & \frac{1}{30} & \frac{1}{6} \\
\frac{1}{7} & \frac{1}{42} & \frac{1}{105} & \frac{1}{140} & \frac{1}{105} & \frac{1}{42} & \frac{1}{7}
\end{array}
$$

人物小传	戈特弗里德·威廉·莱布尼茨(1646—1716)(Gottfried Wilhelm Leibniz)
	微积分的第二个发明人,戈特弗里德·威廉·莱布尼茨(图12.30),生于莱比锡,他的母亲是莱比锡大学哲学系副系主任的第三个妻子.虽然他的父亲在他6岁时就已去世,但年幼的莱布尼茨经过父亲的谆谆教诲,已经产生了读书和学习的愿望.在他年轻时,他就自学了拉丁语并钻研了拉丁文的经典著作以及他父亲丰富藏书中的哲学和神学著作.1661年,他进入莱比锡大学学习,在那里他用大部分时间学习哲学.他在1663年取得学士学位在1664年取得硕士学位,但是尽管准备了法学博士的学位论文,大学却拒绝授给他学位,也许因为教师中的一些政治问题.莱布尼茨因此离开了莱比锡并于1667年从纽伦堡的阿尔特多夫大学取得了学位. 同时,莱布尼茨在1663年在耶纳大学的一次短暂停留中接触到了高等数学,并开始研究他希望是他对哲学最具创造性的贡献的细节问题,创立一种人类思想的字母表,即一种将所有基本概念用符号表示并通过符号的组合表示更复杂的思想的方法.尽管莱布尼茨从未完成这一规划,他的最初的思想包含在他1666年的《论组合的艺术》里,他在论文中独立推导出了帕斯卡的算术三角形以及其中包含的量的各种关系.但这一寻找表达思想的适当符号和组合它们的方式的兴趣最终使他发明了我们今天使用的微积分的符号. 莱布尼茨结束大学学业后不久,他首先为美茵茨选帝侯从事外交方面的工作,而在他以后生涯的大部分时间他是汉诺威公爵的顾问.虽然有许多时期他的工作使他极为忙碌,但他总能找到时间钻研数学思想并在这一领域同遍及欧洲的同事们维持着活跃的通信交流.

调和三角形的每一列通过取第一列与算术三角形相应列的商形成.例如,第3列的元素1/3,1/12,1/30,… 是用帕斯卡三角形第三列的元素3,6,10,… 除以1/1,1/2,1/3,… 得来的.因为每一列包含了它左边的列的元素的差,由此得出每一列中到特定值的元素的和可以根据莱布尼茨的原理通过紧前面一列中最后和第一个元素的差求得.例如$1/2+1/6+1/12=1/1-1/4$.莱布尼茨还注意到这条规则可以推广到无限和,因为取的项数越多,前面一列的最后的值就越小.他因此可以推导出诸如以下的结果

$$\frac{1}{3}+\frac{1}{12}+\frac{1}{30}+\cdots+\frac{1}{n(n+1)(n+2)/2}+\cdots=\frac{1}{2}.$$

给这一序列乘3,莱布尼茨可以将它改写为角锥数的倒数之和:

$$\frac{1}{1}+\frac{1}{4}+\frac{1}{10}+\cdots=\frac{3}{2}.$$



图12.30 德国邮票上的莱布尼茨.

莱布尼茨这里的实际结果并不新.它们的重要性在于当把给差序列求和的思想移植到几何学时所蕴藏的可能性.因而莱布尼茨考虑了定义在分为一些子区间的区间上的一条曲线,并在分割上的每一点x_i建立纵坐标y_i.如果人们构造这些纵坐标差的序列$\{\delta y_i\}$,它的和$\sum_i \delta y_i$等于最末和最开始的纵坐标的差y_n-y_0.同样的,如果人们构造序列$\{\sum y_i\}$,此处$\sum y_i=y_0+y_1+\cdots+y_i$,差的序列$\{\delta\sum y_i\}$等于纵坐标原来的序列.莱布尼茨对这两个规则进行外插来处理有无穷多纵坐标的情况.他把曲线看成是边数为无穷的多边形,在每个交点上向轴作纵坐标y.如果纵坐标的无穷小的差用$\mathrm{d}y$表示,并且无穷多纵坐标的和用$\int y$表示,头一条规则翻译成$\int \mathrm{d}y=y$而第二条规则成为$\mathrm{d}\int y=y$.从几何上看,第一个不过是说一个线段的微分(无穷小的差)的和等于线段.(莱布尼茨在

这里假定初始纵坐标为 0.）第二条规则没有明确的几何解释，因为无穷多有限项的和完全可以是无穷. 所以莱布尼茨将有限的坐标 y 换为无穷小的面积 $y\mathrm{d}x$，此处 $\mathrm{d}x$ 是由无穷多边形的边的交点确定的 x - 轴的无穷小的部分. 因此，$\int y\mathrm{d}x$ 可以被解释为曲线下的面积，而规则 $\mathrm{d}\int y\mathrm{d}x = y\mathrm{d}x$ 不过是说面积 $\int y\mathrm{d}x$ 序列的项的差是项 $y\mathrm{d}x$ 本身.

作为用适当符号表达思想的追求的一部分，莱布尼茨引进了两个记号 d 和 $\int$ 表示他关于差与和的思想的推广. 后者不过是字母拉丁语求和(Summa)的起首字母 S 的加长形式，而前者是拉丁语差(differentia)的起首字母. 对莱布尼茨而言，$\mathrm{d}y$ 和 $\int y$ 的是变量. 换言之，d 和 $\int$ 是赋予有限变量 y 一个无限小变量和一个无穷大变量的算子. 但 $\mathrm{d}y$ 永远被视为变量 y 的两个相邻值的一个实际的差，而 $\int y$ 被看作是从一个固定值到一个给定值变量 y 的实际的和. 因为 $\mathrm{d}y$ 是一个变量，它也能被 d 作用得到一个二阶微分，记作 $\mathrm{dd}y$，或者甚至更高阶的微分. 对一个现代读者来说，想象这些无限小的差和无穷和可能比较困难，但莱布尼茨和他的追随者却在创立许多种问题的解法的过程中对使用这些概念极为擅长.

12.6.2　微分三角形和转换定理

莱布尼茨对微分概念最早的应用之一是在微分三角形的概念上，他从帕斯卡或者还有巴罗那里读到过某种形式的微分三角形. 微分三角形是一个斜边 $\mathrm{d}s$ 连接着表示给定曲线的无穷多边形的两个相邻顶点的无穷小直角三角形，它与由坐标 y、切线 τ 和次切距 t 组成的三角形相似，所以 $\mathrm{d}s : \mathrm{d}y : \mathrm{d}x = \tau : y : t$(图 12.31). 因为切线的概念涉及的是比值，莱布尼茨一般令这三个微分中的一个为常数. 换言之，在选择如何将一个曲线表示为无穷多边形时，他可以或者使多边形有相同的边长($\mathrm{d}s$ 为常数或者 $\mathrm{dd}s = 0$)，使边在 x - 轴上的投影相等($\mathrm{d}x$ 为常数或者 $\mathrm{dd}x = 0$)，或者使边在 y - 轴上的投影相等($\mathrm{d}y$ 为常数或者 $\mathrm{dd}y = 0$). 在某种意义上，被选为具有常数微分的变量可以被视为独立变量. 无论如何，莱布尼茨通过用他操作微分的基本法则对微分三角形中微分的操作，发现了他微积分的核心技巧.

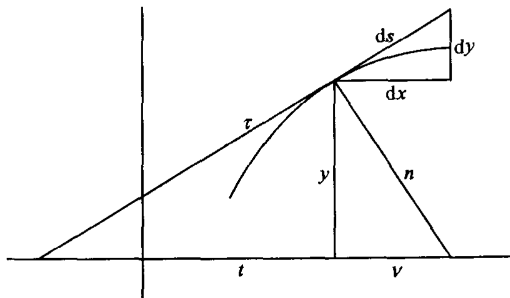


图 12.31　莱布尼茨的微分三角形.

帕斯卡在半径为 r 的圆中利用微分三角形证明了莱布尼茨语言中的 $y\mathrm{d}s = r\mathrm{d}x$. 莱布尼茨意识到因为由纵坐标、法线和次法线 ν 组成的三角形同微分三角形相似，如果人们将半径用法线 n 代替，这一规则可以推广到任意曲线. 所以，$y : \mathrm{d}x = n : \mathrm{d}s$ 或者 $y\mathrm{d}s = n\mathrm{d}x$. 因为 $2\pi y\mathrm{d}s$ 可以被解释为由绕 x - 轴旋转 $\mathrm{d}s$ 形成的曲面的表面积，这一公式将曲面面积的计算替代为面积的计算. 同样，

莱布尼茨注意到 $dx:dy=y:\nu$ 或者 $ydy=\nu dx$. 因为他意识到 $\int ydy$ 代表面积为 $(1/2)b^2$ 的三角形，此处 b 是纵坐标 y 的终值，他得到 $\int \nu dx=(1/2)b^2$. 因此要求纵坐标为 z 的曲线下的面积，只需要找到次法线 ν 等于 z 的曲线 y 就够了. 但因为 $\nu=y(dy/dx)$，这等价于解方程 $y(dy/dx)=z$. 换句话说，一个面积问题可以化归为一个莱布尼茨所称的切线的反问题.

尽管这些特定的规则没有使莱布尼茨得出任何从前尚不为人所知的结果，但这一方法的一个推广使他得到了他的"转换定理"以及圆的算术求积，$\pi/4$ 的一个级数展开式. 在曲线 $OPQD$ 中，此处 P 和 Q 无限靠近，他作三角形 OPQ. 延长 $PQ=ds$ 成为曲线的切线，作 OW 垂直于切线，并且如图 12.32 定义 h 和 z，他利用三角形 TWO 和微分三角形的相似性证明了 $dx:h=ds:z$ 或者 $zdx=hds$. 第二个方程的左边是矩形 $UVRS$ 下的面积，而右边是三角形 OPQ 面积的两倍. 由此得出所有三角形的和，即由曲线 $OPQD$ 和直线 OD 围成的面积，等于纵坐标为 z 的曲线下的面积，或者 $(1/2)\int zdx=\int ydx-(1/2)OG\cdot GD$. 将 OG 表示为 x_0，GD 表示为 y_0，莱布尼茨的转换公式现在可表述为

$$\int ydx=\frac{1}{2}\left(x_0y_0+\int zdx\right).$$

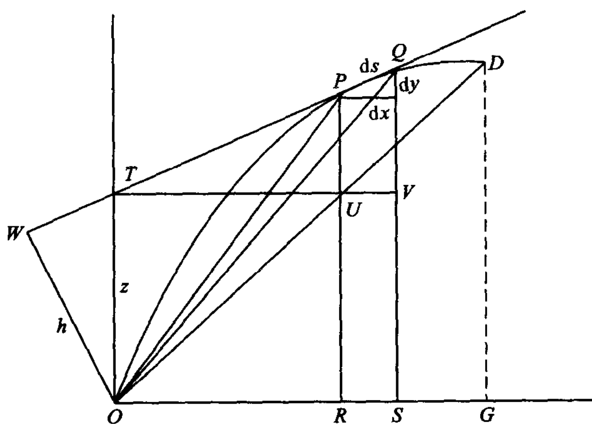


图 12.32　莱布尼茨的转换定理.

因为 $z=y-PU=y-x(dy/dx)$，并且因为莱布尼茨可以用胡德或者斯卢兹的法则计算切线，如果计算 $\int zdx$ 比计算 $\int ydx$ 简单的话，转换定理使他可以计算出原曲线下的面积. 莱布尼茨应用该结果计算了由 $y^2=2x-x^2$ 给出的半径为 1 的四分之一圆的面积. 在这一情况下，

$$z=y-x\left(\frac{1-x}{y}\right)=\frac{x}{y}=\sqrt{\frac{x}{2-x}},$$

或

$$z^2=\frac{x}{2-x},$$

或者，最后 $x=\dfrac{2z^2}{1+z^2}$.

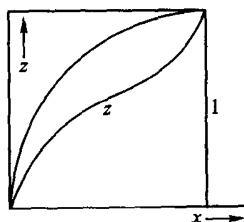


图 12.33　圆的转换函数：$z^2=\dfrac{x}{2-x}$ 或 $x=\dfrac{2z^2}{1+z^2}$.

根据莱布尼茨的转换定理，$\int ydx$（或 $\pi/4$）等于 $(1/2)(1+\int zdx)$. 因为从图 12.33 可以清楚地看

到$\int z\mathrm{d}x = 1 - \int x\mathrm{d}z$，莱布尼茨得出$\int y\mathrm{d}x$等于$1 - \int \frac{z^2}{1+z^2}\mathrm{d}z$的结论.采用和梅卡托类似的论证，他证明了

$$\frac{z^2}{1+z^2} = z^2(1 - z^2 + z^4 - z^6 + \cdots),$$

因此

$$\int y\mathrm{d}x = 1 - \frac{1}{3}z^3 + \frac{1}{5}z^5 - \frac{1}{7}z^7 + \cdots.$$

立即可以得到莱布尼茨算术求积的公式，即 $\pi/4 = 1 - 1/3 + 1/5 - 1/7 + \cdots$.

12.6.3　微分学

莱布尼茨在 1674 年发现了他的转换定理以及圆的算术求积.以后的两年里，他发现了他微分学的全部基本思想.他把这些结果的一部分到 1684 年才以短文的形式在《教师学报》上发表，短文的题目是"适用于有理数和无理数的最大值、最小值和切线的新方法以及它们的一种值得注意的计算方法".在这篇论文中莱布尼茨不情愿地把他的微分 $\mathrm{d}x$ 定义为无穷小量，因为他相信对这些没有严格定义的量会有大量的批评.因此他把 $\mathrm{d}x$ 作为任意有限线段引入.如果 y 是曲线在横坐标为 x 处的纵坐标，并且 τ 是次切距为 t 处曲线的切线，则 $\mathrm{d}y$ 被定义为满足 $\mathrm{d}y : \mathrm{d}x = y : t$ 的直线.他然后陈述了一些基本的运算法则.如果 a 是常数，则 $\mathrm{d}a = 0$；$\mathrm{d}(v \pm y) = \mathrm{d}v \pm \mathrm{d}y$；$\mathrm{d}(vw) = v\mathrm{d}w + w\mathrm{d}v$；以及 $\mathrm{d}(v/y) = (\pm v\mathrm{d}y \mp y\mathrm{d}v)/y^2$.(按照莱布尼茨的意思，商法则中的符号取决于切线的斜率是正还是负.)

莱布尼茨在 1675 年发现了积和商的法则.实际上，他在那年 11 月 11 日的一份手稿中写道："让我们来检验 $\mathrm{d}x\mathrm{d}y$ 是否和 $\mathrm{d}(xy)$ 是一回事以及 $\mathrm{d}x/\mathrm{d}y$ 和 $\mathrm{d}(x/y)$ 是不是一回事."[58] 为检验他的积法则他提出一个例子，例子中 $y = z^2 + bz$ 且 $x = cz + d$.他首先把 $\mathrm{d}y$ 作为 y 值在 $z + \mathrm{d}z$ 和 z 处的差来计算.所以 $\mathrm{d}y = (z + \mathrm{d}z)^2 + b(z + \mathrm{d}z) - z^2 - bz = (2z + b)\mathrm{d}z + (\mathrm{d}z)^2$.因为 $(\mathrm{d}z)^2$ 同 $\mathrm{d}z$ 相比是无穷小，所以他抛弃掉了该项，得到 $\mathrm{d}y = (2z + b)\mathrm{d}z$.同样地，$\mathrm{d}x = c\mathrm{d}z$ 及 $\mathrm{d}x\mathrm{d}y = (2z + b)c(\mathrm{d}z)^2$.他接着写道："但如果你直接计算出 $\mathrm{d}(xy)$ 会得到同样的结果."遗憾的是，莱布尼茨在这里并没有"把它计算出".但是，在手稿靠后的地方，他通过在另一个例子中证明 $\mathrm{d}(x^2)$ 不同于 $(\mathrm{d}x)^2$ 意识到了他的错误.10 天后，他写出了积法则的正确的形式，稍后用差分论证给出了一个简单的证明："$\mathrm{d}(xy)$ 就是两个相邻 xy 的差；令其中之一为 xy，另一个为 $(x + \mathrm{d}x)(y + \mathrm{d}y)$；我们于是得到 $\mathrm{d}(xy) = (x + \mathrm{d}x)(y + \mathrm{d}y) - xy = x\mathrm{d}y + y\mathrm{d}x + \mathrm{d}x\mathrm{d}y$.省略掉量 $\mathrm{d}x\mathrm{d}y$，该项同其余项相比是无穷小，…… 余下 $x\mathrm{d}y + y\mathrm{d}x$."[59] 商法则可以类似地证明.

在 1684 年的论文中，莱布尼茨继续给出了幂法则 $\mathrm{d}(x^n) = nx^{n-1}\mathrm{d}x$ 以及根的法则 $\mathrm{d}\sqrt[b]{x^a} = (a/b)\sqrt[b]{x^{a-b}}\mathrm{d}x$，仍然是没有证明，还注意到假如将根式写成分数幂的话，第一条规律包含了第二条.使用莱布尼茨的记号，链式法则几乎是显然的.例如，要计算 $z = \sqrt{g^2 + y^2}$ 的微分，此处 g 是一个常量，莱布尼茨令 $r = g^2 + y^2$ 并注意到 $\mathrm{d}r = 2y\mathrm{d}y$，并且

$$\mathrm{d}z = \mathrm{d}\sqrt{r} = \mathrm{d}r/2\sqrt{r},$$

将第一个方程代入第二个，他得到

$$\mathrm{d}z = \frac{2y\mathrm{d}y}{2z} = \frac{y\mathrm{d}y}{z}.$$

为说明他新微积分的用处,莱布尼茨讨论了如何确定最小值和最大值.例如他注意到当 v 增加时 dv 将为正,当 v 减少时 dv 将为负,因为 dv 与恒为正的 dx 的比值给出切线的斜率.由此可得,当 v 既不增加也不减少时,$dv = 0$.在该处纵坐标将是最大值(如果曲线下凹)或最小值(如果曲线上凹).那里的切线将是水平的.莱布尼茨进一步注意到,凸性的问题取决于二阶差 ddv:“当纵坐标 v 增加时,它的增量或者差 dv 也增加(就是说,当 dv 为正时,ddv,差的差,也为正;而当 dv 为负时,ddv 也为负),则曲线是上凹的,否则就是下凹的.当增量取最大值或者最小值,或者增量从减小变为增加时,或由增加变为减小时,就有一个拐点,”[60] 即 $ddv = 0$.

在 1684 年论文的最后一个问题中,莱布尼茨给出了一个例子,此例是“应用数学中最困难也是最美妙的问题,离开了我们的微分学或者类似的东西,就没有人能轻易地解决它.”[61] 这就是 1639 年德邦内向笛卡儿提出的求一个次切距等于给定常数 a 的曲线的问题.如果 y 是候选曲线的纵坐标,曲线的微分方程是 $y(dx/dy) = a$ 或者 $a dy = y dx$,莱布尼茨令 dx 为常量,这等价于使横坐标构成一个算术数列.则方程可以写成 $y = k dy$,这里 k 是常量.由此可得,纵坐标 y 同它们的增量 dy 成正比,或者说 y 构成一个几何数列.因为一个 y 的几何数列和 x 的算术数列的关系正像是数与它们的对数的关系,莱布尼茨得出结论说所求的曲线将是一条“对数”曲线.(我们称其为指数曲线,但无论如何,我们的指数曲线和对数曲线是相对于不同坐标轴的相同的曲线.)从莱布尼茨的讨论可以得出,因为 $x = \log y$, $d(\log y) = a(dy/y)$,常量 a 确定用到的特定的对数曲线.

莱布尼茨在 1684 年没有进一步考虑对数曲线,但同约翰·伯努利在若干年后的讨论后,在 1695 年重新回到了对数函数以及指数函数的微分问题.在那年的一篇论文中,他对贝尔纳·纽文铁特(1654 - 1718)的批评作出了回应,该批评说他的方法不足以计算指数表达式 $z = y^x$(x 和 y 都是变量)的微分[62].对微分的直接计算得到 $dz = (y + dy)^{x+dx} - y^x$.应用二项式定理并抛弃掉 dy 的一次以上的项和 $dx dy$ 的积得到方程 $dz = y^{x+dx} + xy^{x+dx-1}dy - y^x$,该微分方程不是齐次的并似乎不能进一步化简,即使是在 $y = b$ 因而 $dy = 0$ 的特殊情形下也是如此.为克服这一困难,莱布尼茨遵照伯努利 1694 年的一个建议,通过给方程 $z = y^x$ 两边取对数从另外的角度来求解问题,得到 $z = x\log y$.则该方程的微分是:

$$a\frac{dz}{z} = xa\frac{dy}{y} + \log y dx,$$

由此可得:

$$dz = \frac{xz}{y}dy + \frac{z\log y}{a}dx \quad 或 \quad d(y^x) = xy^{x-1}dy + \frac{y^x\log y}{a}dx.$$

如果 $x = r$ 是常数,莱布尼茨注意到该法则化归为幂法则 $d(y^r) = ry^{r-1}dy$.

两年后,约翰·伯努利发表了一篇题为“指数微积分的原理”的论文,在论文中他把莱布尼茨的结果加以推广,来求 $y = x^x$, $x^x + x^c = x^y + y$ 以及 $z = x^{y^v}$ 这些方程中的微分关系.他还明确陈述了当 $a = 1$ 时对数微分的标准结果:“对数的微分,不论它们的组成如何,都等于函数的微分除以函数.”[63] 例如,他写道

$$d(\log\sqrt{x^2 + y^2}) = \frac{x dx + y dy}{x^2 + y^2}.$$

12.6.4 基本定理和微分方程

还记得莱布尼茨是从积和差是逆运算的思想开始后来发展成他的微积分的研究的.结果,微积分基本定理完全是显然的.但他在 1680 年前后的一份手稿中扩展了这一思想,他在手稿中首先注

意到，"我用由所有纵坐标和横坐标的差围成的矩形的面积的和来表示图形的面积，" 或记作$\int y\mathrm{d}x$，其次还注意到，我"通过求和或者积分的图形来得到图形的面积；就是求与给定图形的纵坐标的比是和与差的比的纵坐标."[64] 就是说，为求得纵坐标为 y 的曲线下的面积，人们需要求纵坐标为 z 的曲线，满足 $y = \mathrm{d}z$. 莱布尼茨在 1693 年发表在《教师学报》上的一篇论文中使这一思想更为明确，论文中他证明了"求积分的一般问题可以归结为求切线满足给定规律的曲线."[65] 如他所说明的，假如给定曲线的纵坐标是 y，人们可以找到一条曲线 z 使得 $\mathrm{d}z/\mathrm{d}x = y$（一条切线满足给定规律的曲线），则$\int y\mathrm{d}x = z$，或者用现代的符号表示，假定 $z(0) = 0$，

$$\int_0^b y\mathrm{d}x = z(b).$$

但莱布尼茨像牛顿一样对求面积的兴趣不如对解微分方程的兴趣大，特别是在清楚了重要的物理问题可以用这种方程表达了之后. 但他的技巧是不同的. 例如，考虑莱布尼茨在 1693 年讨论过的单位圆中表达弧 y 和它的正弦 x 的关系的方程.[66] 以 $\mathrm{d}y$，$\mathrm{d}t$ 和 $\mathrm{d}x$ 为边的微分三角形同以 $1, x, \sqrt{1-x^2}$ 为对应边的大三角形相似，所以 $\mathrm{d}t = \dfrac{x\mathrm{d}x}{\sqrt{1-x^2}}$（图 12.34）. 根据毕达哥拉斯定理，$\mathrm{d}x^2 + \mathrm{d}t^2 = \mathrm{d}y^2$. 将 $\mathrm{d}t$ 的值代入并化简，莱布尼茨得到了联系弧长和正弦的微分方程：$\mathrm{d}x^2 + x^2\mathrm{d}y^2 = \mathrm{d}y^2$. 将 $\mathrm{d}y$ 考虑成常量，他对该方程应用他的算子 d 得到 $\mathrm{d}(\mathrm{d}x^2 + x^2\mathrm{d}y^2) = 0$ 或者应用积的法则，$2\mathrm{d}x(\mathrm{dd}x) + 2x\mathrm{d}x\mathrm{d}y^2 = 0$. 莱布尼茨将它简化成二阶微分方程

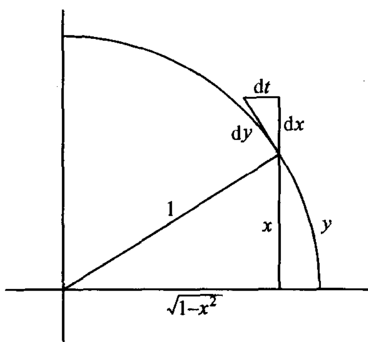


图 12.34　莱伯尼茨对正弦的微分方程的推导.

$$\mathrm{d}^2x + x\mathrm{d}y^2 = 0 \quad \text{或} \quad \frac{\mathrm{d}^2x}{\mathrm{d}y^2} = -x,$$

熟悉的正弦的微分方程.（注意莱布尼茨运用二阶微分的方法解释了我们二阶导数的现代记号中似乎奇怪的对 2 的放置.）

给定了微分方程，莱布尼茨接着假定 x 可以写成 y 的幂级数：

$$x = by + cy^3 + ey^5 + fy^7 + gy^9 + \cdots,$$

其中的系数待定. 对他来说显然不存在偶数次的项并且因为 $\sin 0 = 0$，常数项也为 0. 将这一级数微分两次得到$\dfrac{\mathrm{d}^2x}{\mathrm{d}y^2} = 2\cdot 3cy + 4\cdot 5ey^3 + 6\cdot 7fy^5 + 8\cdot 9gy^7 + \cdots$，这一级数将等于表示 $-x$ 的幂级数. 系数的等同给出一系列简单方程：

$$2\cdot 3c = -b,$$
$$4\cdot 5e = -c,$$
$$6\cdot 7f = -e,$$
$$8\cdot 9g = -f,$$
$$\cdots = .$$

令 $b = 1$ 作为二次初始条件，莱布尼茨轻易地解出了这些方程得到

$$c = -\frac{1}{3!}, e = \frac{1}{5!}, f = -\frac{1}{7!}, g = \frac{1}{9!}, \cdots.$$

由此得到正弦级数，他于 1676 年发现该级数：

$$x = \sin y = y - \frac{1}{3!}y^3 + \frac{1}{5!}y^5 - \frac{1}{7!}y^7 + \frac{1}{9!}y^9 + \cdots.$$

17世纪早期,莱布尼茨发现了当今微积分课本中表述的绝大多数思想,但他从未写出一个对这些材料的完整的一致的表述.无论如何,他像牛顿一样在某种程度上被他对无穷小量的使用所困扰.所以,在他后期的一些论著中,他试图说明使用无穷小量的正当性.虽然他从未说明他是否相信无穷小量实际存在,他通过两种不同的路线说明了其过程的正当性.首先,他试图将无穷小量同阿基米德的穷竭法联系起来:"因为人们把量取得如此大或者小以至于误差可以小于给定的误差,而不是取得无穷大或者无穷小,因此人们同阿基米德的风格仅在形式上不同,我们的方法更直接,更符合发明的艺术."[67] 因而,他似乎像开普勒一样地认为任何使用无穷小的论证都可以按照希腊人的风格用完全严格的论证所代替.但假如人们永远都必须给出这些论证,人们将永远无法获得新的洞察力.莱布尼茨的第二条路线是使用连续的规律:"如果任何连续的变换将在某个极限终止,则有可能构造一个也包含最后极限在内的一般的推理."[68] 换言之,如果人们确定一个特定的比值当量——举例说——dx 和 dy 是有限的时候普遍成立,则同样的比在当这些量自身等于0的极限情况下也成立.无论如何,无穷小量是否存在的问题对莱布尼茨成了无足轻重的问题.基本的计算给出了正确的结果并且所有的无穷小量在最后都被比值和无穷小量相等的有限的量所取代.但是,处理这些无穷小量的技巧变得极为有用,特别是对莱布尼茨的直接后继者约翰·伯努利和雅各布·伯努利更是这样.他们似乎接受无穷小量为实际的数学实体,并且他们通过对无穷小量的运用在微积分本身以及微积分在物理问题中的应用方面都取得了许多重要的成果.

在这里稍微谈谈莱布尼茨和牛顿之间关于优先权的争论是适当的[69].应当弄清楚的是虽然实际上发现的是现在被通称为微积分的同样的规则和过程,他们对该学科的研究是截然不同的.牛顿的研究是从速度和距离的概念着手的,而莱布尼茨则是通过差与和开展研究的.不过,因为牛顿的工作直到18世纪早期才发表,虽然在早得多的时候它们就在英国为人所知了,莱布尼茨和伯努利兄弟在应用他们的版本上的成功使得某些英国数学家指控莱布尼茨剽窃,特别是因为莱布尼茨在17世纪70年代对伦敦的短暂访问中曾经读到过一些牛顿的材料,并且他还通过皇家学会的秘书亨利·奥尔登堡收到过牛顿的两封信,牛顿在信中讨论了他的一些结果.反过来,正是因为牛顿没有发表,伯努利兄弟指控牛顿剽窃了莱布尼茨.1711年,牛顿当时任会长的皇家学会指派了一个委员会调查这些指控.自然,委员会发现莱布尼茨犯有指控中所说的剽窃.争论的不幸结局是英国和大陆的数学家实际上停止了思想交流.就微积分来说,英国的数学家采用了牛顿的方法和符号,而在大陆上,数学家们采用了莱布尼茨的方法和符号.结果表明,莱布尼茨的符号和他的微分学更容易使用,因此分析学的进步在大陆出现得更快.最大的损失是,英国数学团体在几乎整个18世纪里剥夺了自己取得显著进步的机会.

12.7 第一批微积分教科书

英国和大陆之间研究路线的差别生动地表现在第一批微积分教科书中,这批教科书分别是由法国的洛必达侯爵(Guillaume Francois l'Hospital,1661—1704)在1696年、英国的查里斯·海耶斯(Charles Hayes,1678—1760)和哈弗里·狄顿(Humphry Ditton,1675—1715)在1704年和1706年完成的.本章的最后我们将对这些早期教科书的某些方面作一简要研究以便使读者们对早期微积分的学生需要掌握的东西有一概念.

12.7.1　洛必达的《无穷小量分析》

洛必达出生在一个贵族家庭，年轻时代曾作过军官. 大约在 1690 年，他开始对新分析产生兴趣，当时杂志上才刚开始出现莱布尼茨和伯努利兄弟的有关文章. 不幸的是，这些文章经常简洁得有些晦涩，至少在方法方面如此. 因为约翰 · 伯努利 1691 年时在巴黎逗留，洛必达支付优厚的报酬请他讲授那个新的学科. 伯努利同意了并且讲解了一部分. 大约 1 年后，伯努利离开巴黎到荷兰的格罗宁根大学去作教授. 因为洛必达希望讲授能继续下去，他们在一笔丰厚月薪的基础上达成协议，伯努利不仅会继续寄给洛必达微积分方面的材料，包括他自己可能作出的任何新发现，而且还将拒绝透露给别人. 实际上，伯努利在为洛必达工作. 到了 1696 年，洛必达确信他已经懂得了足够多的微积分，可以发表一本关于它的教科书了，因为他觉得他已经对伯努利的工作支付了不错的报酬，他便没有愧疚地使用了大量后者在新数学中的组织和发现. 虽然伯努利对他的工作被别人发表有些不悦，出版者对他仅仅作了致谢，他对事情保持了沉默. 因为洛必达在一本积分学的著作出版前就已去世，伯努利最终在这方面发表了他自己的讲义.

洛必达在这本最早出现但极为成功的题为《理解曲线的无穷小量分析》的微积分教科书中，首先将变量定义为那些持续增加或减小的量，由此给出他对微分的基本定义："一个变量持续增加或减小的无穷小的部分称为该变量的微分." 然后他提出了规范他对这些微分的使用的两条公理：

1. 承认两个相差无穷小量的量可以被彼此不加区分的理解（或使用）；或者（还是同一件事情）一个只增加或减小了无穷小量的量可以被当作保持不变.

2. 承认一条曲线可以被理解为由无数无限小的直线组成；或者（还是同一件事情）被理解为一个边数为无限的多边形，每条边无穷小，通过它们之间的夹角确定曲线的曲率.[70]

对当时的洛必达来说，无穷小的存在性不是问题. 它们存在，它们能够用微分三角形的元素表示，并用他提出的各种法则加以运算. 法则的表述和证明基本上以莱布尼茨最初使用的同样的方式进行，但虽然莱布尼茨和伯努利兄弟已经开始在他们的工作中考虑超越曲线了，洛必达实际上单单处理代数曲线. 他仅仅提及对数曲线，将其定义为次切距 $y(\mathrm{d}x/\mathrm{d}y)$ 是常量的曲线，而且根本没有涉及任何类似三角曲线的东西.

洛必达对最大值和最小值的处理比莱布尼茨的处理稍微普遍一些. 他注意到如果纵坐标增加则微分 $\mathrm{d}y$ 为正；如果纵坐标减小则为负，但进一步说明了有两种方式 $\mathrm{d}y$ 可以从正变为负，而纵坐标从增加变为减小，即，当 $\mathrm{d}y$ 越过 0 或者越过无穷时. 作为这种讨论的一部分，他用图说明了 4 种可能性，两个有水平的切线，两个有尖点而切线是竖直的，他还用例子说明了这些可能性. 例如，为计算 $y - a = a^{1/3}(a-x)^{2/3}$，他计算

$$\mathrm{d}y = -\frac{2\sqrt[3]{a}\,\mathrm{d}x}{3\sqrt[3]{a-x}}.$$

因为 $\mathrm{d}y = 0$ 是不可能的，他令 $\mathrm{d}y$ 等于无穷大. 这意味着 $3\sqrt[3]{a-x} = 0$ 或者 $x = a$. 洛必达没有给出区分最大值和最小值的特别的方法，但极值的性质通常从问题的条件可以弄清楚. 例如，考虑求具有给定体积 a^3 且一边等于定长 b 的平行六面体使其表面积为最小的现在的标准问题. 因为平行六面体的边为 b，x 和 a^3/bx，问题归结为求 $y = bx + a^3/x + a^3/b$ 的最小值的问题. 洛必达的结论是最小值出现在 $x = \sqrt{a^3/b}$ 这一点.

洛必达自然讨论了二阶微分并像莱布尼茨一样得出拐点出现在 $\mathrm{d}^2y = 0$ 一点，假定取 $\mathrm{d}x$ 为常

量.使用了同牛顿类似的方法,他也推导出了确定给定曲线曲率半径的公式

$$r = \frac{(dx^2 + dy^2)^{3/2}}{dx d^2 y}.$$

但洛必达的《分析》可能作为在分子和分母的极限都为0时计算商的极限用到的洛必达法则的来源时最为出名 —— 该法则也许应该改称伯努利法则.

命题 设 AMD 是一条曲线($AP = x, PM = y, AB = a$)满足纵坐标 y 的值用一个分数表示,该分数当 $x = a$ 时分子和分母都变为0,就是说,当点 P 与给定的点 B 重合时.则需要求的便是纵坐标 BD 的值(图 12.35).[71]

图 12.35 洛必达对洛必达法则的图解.注意,函数 q 被画在 x 轴的下方,而商函数在 x 轴的上方,即 AMD 曲线所表示的.把所涉及的函数值均想像成为正的量.

设 $y = p/q$,洛必达简单地注意到对无限接近 B 的横坐标 b,纵坐标 y 的值将由下式给出

$$y + dy = \frac{p + dp}{q + dq}.$$

但因为该纵坐标无限趋近 y,并且在点 B 处 p 和 q 均为0,洛必达简单地使 y 等于 dp/dq.换言之,"如果求得了分子的微分,用它除以分母的微分,令 $x = a$ 后…,我们将得到要求的纵坐标的值."(注意在这一定理的表述和证明中没有涉及到极限.)洛必达在这里不相信平凡的例子.他的第一个例子是函数

$$y = \frac{\sqrt{2a^3x - x^4} - a\sqrt[3]{a^2x}}{a - \sqrt[4]{ax^3}},$$

该例子是若干年前伯努利告诉他的,当 $x = a$ 时可以求得纵坐标的值.他通过对微分直截了当的计算得到 $y = (16/9)a$.

12.7.2 狄顿和海耶斯的著作

转过来看英国作家,我们在狄顿的《一种流数体系》和海耶斯的《论流数》中找到了一种有些不同的课本.这两个作者并不是很有名.但他们都学习了新的微积分,并且给他们的学生讲授,他们感觉到一种英语课本对学生是有用的.虽然他们都读过大陆作家的著作,但他们自然更偏向于牛顿的流数研究而不是莱布尼茨的微分研究.因此狄顿写道数量不会被想象成"无穷多小的组成元素的集合或者和而是被想象成持续变动的结果,变动不断地进行,从开始的第一刻一直到完全静止的一刻.直线不被描述成小线段或部分的并列,而是被描述成点的连续运动,…… 作为流数法基础的基

本原理比微积分的基本原理更准确、清楚并使人信服."[72] 狄顿接着试图使读者相信微分学中因为某些项是"没有"而将其抛弃掉的作法不如流数论中因为某些项与"最终真正消逝"的量相乘而将其舍弃的作法正当.不论这些哲学的论证是否能使学生们信服,海耶斯和狄顿都像牛顿一样对微积分学的两个分支的基本内容给出了清楚的阐述.

只有对对数函数和指数函数的精细微积分方面,这可能是从伯努利的论文中学到的,以及对积分学的处理上,这些书同洛必达著作的内容有所不同.例如,两人都证明了伯努利定理,任何量的对数的流数等于该量的流数除以该量.(用流数的符号表示是 $\dot{\ell}(x) = \dot{x}/x$,式中 $\ell(x)$ 表示对数.)狄顿用幂级数给出了证明:因为 $\ell(1+x) = x - x^2/2 + x^3/3 - x^4/4 + \cdots$,由此推出 $\dot{\ell}(1+x) = \dot{x} - x\dot{x} + x^2\dot{x} - x^3\dot{x} + \cdots = \dot{x}(1 - x + x^2 - x^3 + \cdots) = \dot{x}/(1+x)$,该结果同要求的定理等价.

考察了对数之后,两位作者转向了指数函数 $y = a^x$,他们用从莱布尼茨的过程直接翻译过来的方法处理它们.为计算 y 的流数,作者注意到 $\ell(y) = x\ell(a)$,并且两边取流数,计算出 $\dot{y}/y = \dot{x}\ell(a) + x\dot{\ell}(a)$.因为 a 是常量,它的对数的流数为 0.由此得出 $\dot{y} = y\dot{x}\ell(a) = a^x\ell(a)\dot{x}$.

海耶斯还考虑了由指数函数确定的曲线.还记得对数曲线是当横坐标成算术数列时纵坐标成几何数列的曲线或者说次切距为常数的曲线.因为任何曲线 y 的次切距由 $y(\dot{x}/\dot{y})$ 给出,并且因为曲线 $y = a^x$ 的次切距由下式给出

$$y\frac{\dot{x}}{y\dot{x}\ell(a)} = \frac{1}{\ell(a)},$$

由 $y = a^x$ 定义的曲线一定是对数曲线.进而,海耶斯通过首先注意到面积的流数一般是 $y\dot{x}$ 计算了对数曲线下的面积.因为在这一情况下,曲线的次切距 $y(\dot{x}/\dot{y})$ 是常量 c,由此得出 $y\dot{x} = c\dot{y}$ 并且面积的流数因此是 $c\dot{y}$ 而面积本身必定是 cy.海耶斯的结论是任何两个横坐标间的对数曲线下的面积与相应纵坐标的差成正比,这一结果在莱布尼茨和伯努利的著作中都没有明确给出.

狄顿详细处理了积分学的其它方面,包括求曲线长度、求曲面面积、求立体体积以及重心.但他的课本和海耶斯或者洛必达的课本一样没有包括对正弦或者余弦的微积分的处理.作为某些问题的一部分对这些三角关系偶有提及,但在 18 世纪 17 世纪之交还没有出现对这些函数的微积分的处理.它们直到 18 世纪 30 年代在莱奥纳德·欧拉的著作中才出现.

习　题

关于切线和极值的问题

1. 说明能够内接于球内的最大的平行六面体是立方体.当球的半径为 10 时确定立方体的尺寸和体积.
2. 说明能够内接于球内的最大的圆柱体的直径与高度的比值是$\sqrt{2}:1$.(开普勒)
3. 说明费马确定多项式 $p(x)$ 最大值和最小值的两种方法都等价于解 $p'(x) = 0$.
4. 用费马的一种方法求 $bx - x^3$ 的最大值.费马如何确定选择两个解中的哪一个作为他的最大值?
5. 说明如果 M 是$p(x)$ 的最大值,则多项式 $p(x) - M$ 总有因式$(x-a)^2$,其中 a 是给出最大值的 x 的值,借此说明费马确定最大值和最小值的第一种方法的正当性.
6. 使用费马的切线法确定点 B 的横坐标 x 和给出 $y = x^3$ 切线的次切距 t 间的关系.
7. 对费马的切线法加以修正使之能够适用于形如 $f(x,y) = c$ 的方程确定的曲线.首先须注意到如果$(x+e,\bar{y})$是切线上靠近(x,y)的一点,则 $\bar{y} = \frac{t+e}{t}y$.然后将$f(x,y)$和$f\left(x+e,\frac{t+e}{t}y\right)$等同.运用该方法确定曲线 $x^3 + y^3$

$= pxy$ 的次切距.

8. 说明在现代的符号中,费马求曲线 $y = f(x)$ 次切距 t 的方法用式 $t = f(x)/f'(x)$ 确定 t. 用类似的方法说明在练习 7 中修正过的方法在现代的术语中等价于用式 $t = -y(\partial f/\partial y)/(\partial f/\partial x)$ 来确定 t.
9. 用费马的方法确定椭圆 $x^2/a^2 + y^2/b^2 = 1$ 的次切距. 将你的答案同第三章中阿波罗尼乌斯的答案作比较.
10. 用笛卡儿的圆法确定 $y = x^{3/2}$ 次法线.
11. 用笛卡儿的圆法确定 $y^2 = x$ 的切线的斜率.
12. 在笛卡儿的方法中用胡德法则说明 $y = x^n$ 在 (x_0, x_0^n) 处的切线的斜率是 nx_0^{n-1}.
13. 用胡德法则将 $3ax^3 - bx^3 - \frac{2b^2a}{3c}x + a^2b$ 最大化.(该例子系从胡德的《论最大值和最小值》中选出).
14. 对费马的方程 $x^3 + y^3 - pxy = 0$ 应用斯卢兹法则.
15. 用斯卢兹法则求圆 $x^2 + y^2 = bx$ 的切线.
16. 用费马确定 $y = g(x)$ 次切距的法则在 $f(x,y) = g(x) - y$ 的特殊情形下推导出斯卢兹法则. 从练习 7 中修正过的费马法则中推导出普遍的斯卢兹法则.

关于面积和体积的问题

17. 已知圆锥的体积为 $(1/3)hA$,其中 h 是高而 A 是底面积,用开普勒的方法将半径为 r 的球分成无穷多个无穷小的高为 r 的圆锥并把它们的体积加总得到球的体积公式.
18. 证明费马的法则

$$N\binom{N+k}{k} = (k+1)\binom{N+k}{k+1}$$

等价于

$$N\sum_{j=k-1}^{N+k-1}\binom{j}{k-1} = (k+1)\,\frac{N(N+1)\cdots(N+k)}{(k+1)!},$$

以及

$$\sum_{j=1}^{N}\frac{j(j+1)\cdots(j+k-1)}{k!} = \frac{N(N+1)\cdots(N+k)}{(k+1)!}.$$

19. 在练习 18 的最后一个公式中令 $k = 3$ 并从这一结果和已知的整数求和公式和平方和公式导出立方和公式.
20. 用练习 18 中的公式找出求 $\sum_{j=1}^{N} j^4$ 的 5 次多项式公式.
21. 用对 k 归纳的方法证明

$$\sum_{j=1}^{N} j^k = \frac{N^{k+1}}{k+1} + \frac{N^k}{2} + p_{k-1}(N),$$

式中 $p_{k-1}(N)$ 是次数低于 k 的多项式.

22. 费马在 1636 年 8 月 23 日给罗伯华的一封信中有以下结果:如果一个顶点为 A,轴为 AD 的抛物线绕直线 BD 旋转,所得立体的体积与具有相同顶点和底的圆锥的体积之比为 8 : 5(图 12.36). 证明费马是正确的并说明该结果等价于伊本 · 海塞姆发现的并在第七章中讨论过的计算相同立体的体积的结果.
23. 通过把区间 $[0, x_0]$ 分为子区间的无穷集合的方法来确定从 $x = 0$ 到 $x = x_0$ 的曲线 $y = px^k$ 下的面积,子区间从右开始于点 $a_0 = x_0, a_1 = (n/m)x_0, a_2 = (n/m)^2x_0, \cdots$,式中 $n < m$,然后按照费马推导双曲线下面积的方法进行.
24. 用无穷小量方法说明沃利斯关于和的比的结果意

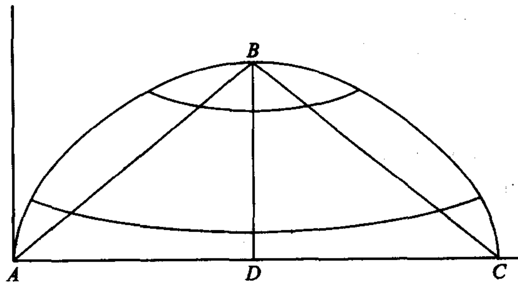


图 12.36　绕垂直于其轴的直线旋转的抛物线的费马问题.

味着

$$\int_0^{x_0} x^{p/q}\mathrm{d}x = \frac{q}{p+q}x_0^{p+q/q}.$$

25. 用沃利斯的方法为他的比率表中插入对应着 $p = 3/2$ 和 $n = 3/2$ 的行和列.

26. 通过使用 Viscount Brounker William(1620—1684) 的将面积分成矩形的方法逼近介于 $x = 0$ 和 $x = 1$ 间双曲线 $y = \frac{1}{x+1}$ 下的面积,以此计算出 log 2 的无穷级数(图 12.37).说明

$$R_1 = \frac{1}{1\times 2}, R_2 = \frac{1}{3\times 4}, R_3 = \frac{1}{5\times 6}, R_4 = \frac{1}{7\times 8}.$$

然后找出这一级数的下面 4 项.最后得出该级数继续的一般法则并推导出

$$\log 2 = \frac{1}{1\times 2} + \frac{1}{3\times 4} + \frac{1}{5\times 6} + \frac{1}{7\times 8} + \cdots.$$

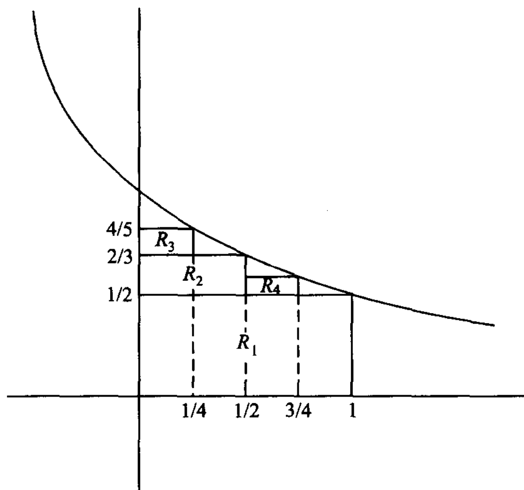


图 12.37　用布隆克尔的方法确定双曲线 $y = \frac{1}{x+1}$ 下方的面积.

格利高里、范·休莱特和巴罗的问题

27. 用泰勒公式推导出正切函数的幂级数(到第 9 次幂的项)并把它同格利高里的结果加以比较.格利高里出错了吗?

28. 格利高里推导出了计算由其它曲线通过加、减、比例复合成的曲线的次切距的各种公式并用这些公式计算幂级数.特别的,假定 4 个函数通过比例关系 $u : v = w : z$ 相关联.说明次切距 t_z 由下列公式给出:

$$t_z = \frac{t_u t_v t_w}{t_u t_v + t_u t_w - t_v t_w}.$$

从该公式推导出导数的积法则和商法则,已知如果函数 u 是常量,它的次切距 t_u 是无穷大.

29. 求曲线 $y^4 = x^5$ 从 $x = 0$ 到 $x = b$ 的弧长.

30. 说明为求抛物线 $y = x^2$ 的弧长人们需要求双曲线 $y^2 - 4x^2 = 1$ 下的面积.

31. 用巴罗的 a,e 法求曲线 $x^3 + y^3 = c^3$ 的切线的斜率.

32. 巴罗可能是使用他的 a,e 法计算曲线 $y = \tan x$ 的切线斜率的第一个人.假定 DEB 是单位圆的一个象限,BX 是 B 处的切线(图 12.38).切曲线 AMO 被定义为满足以下条件的曲线:如果 AP 等于弧 BE,则 PM 等于弧 BE 的切线 BG.用微分三角形如下计算曲线 AMO 的斜率:令 $CK = f$ 且 $KE = g$.因为 $CE : EK =$ 弧 $EF : LK$,由此可得 $1 : g = e : LK$ 或者 $LK = ge$ 以及 $CL = f + ge$.则

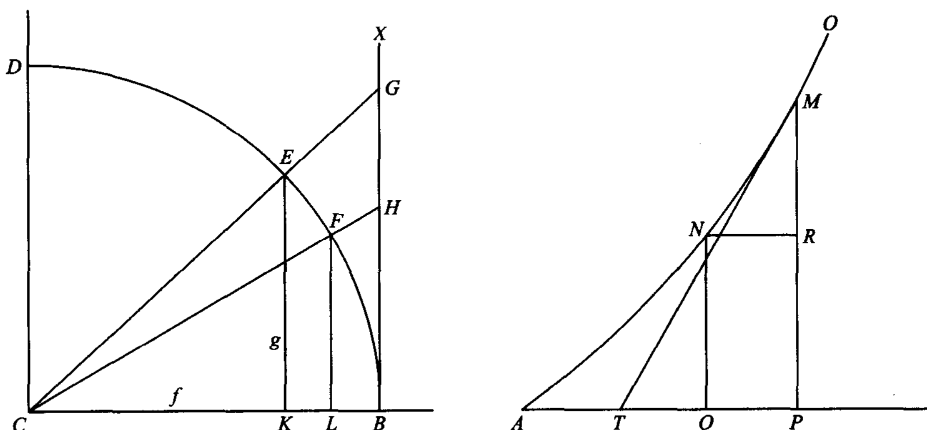


图 12.38　巴罗对正切函数的切线的计算.

$$LF = \sqrt{1 - f^2 - 2fge} = \sqrt{g^2 - 2fge}.$$

因为 $CL : LF = CB : BH$,人们就可以将圆中的比例转移到切曲线上.最后说明

$$PT = t = -\frac{BG \cdot CB^2}{CG^2} = \frac{BG \cdot CK^2}{CE^2}.$$

并说明该结果可以被翻译成大家熟悉的公式 $\mathrm{d}(\tan x)/\mathrm{d}x = \sec^2 x$.承认该结果,人们能否说巴罗给出了三角函数的微分?为什么或者为什么不?

牛顿的问题

33. 用对 $1 + x$ 开平方的算法计算 $\sqrt{1 + x}$ 的幂级数.
34. 用长除法计算 $1/(1 - x^2)$ 的幂级数.
35. 用牛顿的方法解方程 $x^2 - 2 = 0$ 使结果精确到小数点后8位.这需要多少步?将该方法的效率同中国开方算法加以比较.
36. 解 $y^3 + y - 2 + xy - x^3 = 0$ 中的 y 为 x 的幂级数.从解 $x = 0$ 时的 y 的值开始,即先解 $y^3 + y - 2 = 0$.因为 $y = 1$ 是一个解,假定 $y = 1 + p$ 是原方程的解.将这一 y 值代入得到 $1 + 3p + 3p^2 + p^3 + 1 + p - 2 + x + px - x^3 = 0$.删去所有 x 和 p 高于1次的项,解 $4p + x = 0$ 得到 $p = -(1/4)x$.因此 $1 - (1/4)x$ 是所求 y 的级数的前两项.进而,将 $p = -(1/4)x + q$ 代入 p 的方程并继续原来的技巧.说明该级数的下一项是 $(1/64)x^2$.
37. 用练习 36 中的牛顿法解方程 $(1/5)y^5 - (1/4)y^4 + (1/3)y^3 - (1/2)y^2 + y - z = 0$ 的 y.从第一近似 $y = z$ 开始.然后把 $y = z + p$ 代入级数,删掉 p 的非线性项并解得 $y = z + (1/2)z^2$ 作为第二近似.如此继续再得到该级数的两项,$(1/6)z^3$ 和 $(1/24)z^4$.
38. 将 $(1 - x^2)^{1/2}$ 的牛顿级数平方并说明所得的幂级数等于 $1 - x^2$.(你需要使自己相 $x2$ 以上的项的所有系数都等于0.)
39. 用 $\log(1 + x)$ 的幂级数计算 $1 \pm 0.1, 1 \pm 0.2, 1 \pm 0.01, 1 \pm 0.02$ 的对数到小数点后8位.用本书中给出的恒等式和你自己构造出的其它恒等式计算出精确到小数点后8位的从1到10的整数的对数表.
40. 用数列4,3,2,1作乘数计算方程 $x^3 - ax^2 + axy - y^3 = 0$ 中流数的关系.你注意到了什么?如果你用了一个不同的数列会是什么结果?
41. 对方程 $y^2 - a^2 - x\sqrt{a^2 - x^2} = 0$ 用牛顿法则计算流数的关系.设 $z = x\sqrt{a^2 - x^2}$.

42. 首先用 $x+1$ 代换 x 再使用级数技巧解流数方程 $\dot{y}/\dot{x} = 2/x + 3 - x^2$.
43. 假定在一个简化的太阳系中所有的行星在以太阳为中心的圆上作匀速运动.如果向心力同半径的平方成反比,说明行星周期的平方等于半径的立方.(这是开普勒第三定律的一个特殊情形.)

莱布尼茨的问题

44. 从调和数列 1/1,1/2,1/3,1/4,…… 开始并取差的方法构造莱布尼茨的调和三角形.建立三角形中元素的一个公式.
45. 给定曲线 $y^q = x^p(q > p > 0)$,用转换定理证明$\int_0^{x_0} y\mathrm{d}x = \frac{qx_0y_0}{p+q}$.注意从 $y^q = x^p$ 可以得到

$$\frac{q\mathrm{d}y}{y} = \frac{p\mathrm{d}x}{x},\text{且因此 } z = y - x\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{q-p}{q}y.$$

46. 使用微分论证证明商法则 $\mathrm{d}\left(\frac{x}{y}\right) = \frac{y\mathrm{d}x - x\mathrm{d}y}{y^2}$.
47. 从微分方程 $\mathrm{d}y = \frac{1}{x+1}\mathrm{d}x$ 推导出对数的幂级数,假设 y 是 x 系数未定的幂级数,求解简单的方程来依次确定各个系数.
48. 已知 $x+1$ 的对数 y,用待定系数法推导确定出数 $x+1$ 的幂级数,这正是莱布尼茨所说的指数函数的幂级数.从微分方程 $x + 1 = \mathrm{d}x/\mathrm{d}y$.

早期微积分教科书中的问题

49. 对洛必达本人的例子 $y = \frac{\sqrt{2a^3x - x^4} - a\sqrt[3]{a^2x}}{a - \sqrt[4]{ax^3}}$ 用洛必达法则求当 $x = a$ 时 y 的值.
50. 用海耶斯和狄顿的方法计算 $y = x^x$ 的流数.

讨 论

51. 从它们的符号、使用的方便程度和它们的基础方面比较牛顿和莱布尼茨的“微积分学”的相同点和不同点.
52. 比较费马的切线法和笛卡儿的圆法在确定曲线 $y = x^n$ 的切线的斜率上的效率.注意在每例中需要作的运算的种类.
53. 规划一堂介绍积分概念的课,将费马的方法应用到方程为 $y = x^n$ 类型的曲线,其中 n 是正整数.
54. 用牛顿的思想方法规划一系列关于幂级数的课.在微积分的课程早早引入这些级数有用吗?为什么或者为什么不?
55. 能否沿着牛顿《论方法》的路线建立微积分课程的框架?这同通常的微积分课程的组织有什么不同?如何将这种组织同微积分课程的一种改革方案的材料组织加以比较?
56. 用范·休莱特的方法规划一堂介绍弧长确定的课.这同在微积分课本中通常采用的方法有什么不同?
57. 在当今的微积分课堂 —— 传统的或者革新的 —— 上,作为无穷小量的微分概念有用吗?它能否使微积分基本法则的推导更为容易?为什么或者为什么不?
58. 遵循牛顿类比论证的方法,规划一堂关于广义二项式定理的课.
59. 为什么把牛顿和莱布尼茨而不是在本章中考虑过的一些更早的数学家当成是微积分的发明人?

文献和注解

在微积分历史方面有三本当代的著作,它们共同构成了对有关材料的相当完整的论述.最早的一本是 Carl

Boyer 的 The History of the Calculus and its Conceptual Developmen(New York:Dover,1959).这部著作主要论述作为微积分基础的核心概念.虽然它通常仅考虑那些预示着近代发展的概念,而不是将这些概念作为它们所处时代的一部分加以考虑,但该著作仍然是包含了在本章以及同微积分有关的前面和后面几章研究过的大多数概念的一个绝佳论述.Margaret E. Baron 的 The Origins of the Infinitesimal Calculus(Oxford:Pergamon Press,1969)更多地追踪了直到牛顿和莱布尼茨时代在微积分中实际使用的方法.特别是该书在论述 17 世纪的第一部分非常出色,它提供了许多有助于我们理解各个数学家实际中如何解决他们碰到的问题的事例.第三本书是 C.H.Edwards 的 The Historical Development of the Calculus (New York:Springer,1979),它也致力于精确讲述数学家是如何计算的,但不像前一本书,它不仅详细论述了牛顿和莱布尼茨的贡献,还论述了他们在 18、19 世纪的后继者的工作.对微积分史的一个简明综述可以在 Arthur Rosenthal 的"The History of Calculus", American Mathematical Monthly 58(1951),75 - 86 中找到.对 17 世纪数学的更一般的讨论可以在 D.T.Whiteside 的"Patterns of Mathematical Thought in the Later Seventeenth Century"Archive for History of Exact Sciences 1(1960 - 62),179 - 388 中找到.微积分思想在 20 世纪早期发展的进一步的信息可以在 Ivor Grattan-Guinness 主编的 From the Calculus to Set Theory, 1630 - 1910: An Introductory History(London:Duckworth,1980).

1. J.M.Child, *The Early Mathematical Manuscripts of Leibniz*(Chicago:Open Court,1920),p.215.这部著作收录了许多经过编辑的莱布尼茨直到大约 1680 年的著作的译文.书中的评论必须被谨慎地阅读,因为 Child 似乎竭力要说明莱布尼茨的许多工作可以从伊萨克·巴罗的工作中导出.
2. H.W.Turnbull 主编 *The Correspondence of Isaac Newton*(Cambridge:Cambridge University Press,1960),vol.Ⅱ,p.115, 153.这一七卷本的集子实际上收录了现存的所有给牛顿或从牛顿处收到的信件,以及其它有关的一些材料.在翻译那段密文时,必须记住在拉丁文中,字母 u 和 v 是可以互换的.
3. M.Caspar 主编,Johann Kepler, *Gesmmelte Werke*,(Munich:Beck,1960),vol.Ⅳ,p.85.
4. 转引自 Michael Mahoney, *The Mathematical Career of Pierre de Fermat*,1601—1665(Princeton:Princeton University Press, 1973),p.148.
5. D.J.Struik 主编,*A Source Book in Mathematics*,1200—1800(Cambridge:Havard University Press,1969),p.223.
6. 有关费马和笛卡儿间围绕费马方法的争论以及梅森在中间起到的煽风点火的作用的讨论,见 Mahoney, *The Mathematical Career*,第九章.
7. David Eugene Smith 和 Marcia L.Latham 合译的 *The Geometry of René Descartes*(New York:Dover,1954),p.95.
8. 同上,pp.100 - 104.
9. Struik,Source Book,p.194.
10. 同上,p.196.
11. 转引自 Edwards,*Historical Development*,p.103.
12. 转引自 Francois De Gandt 的 *Force and Geometry in Newton's Principia*(Princeton:Princeton University Press,1995),pp. 172 - 173.该书探索了作为牛顿 1687 年代表作基础的一些基本的数学和物理学思想.
13. 同上,p.176.
14. Kirsti Andersen,"Cavalieri's Method of Indivisibles",*Archive for History of Exact Sciences* 31(1985),291 - 367,p.316.对卡瓦列里的讨论主要以这篇论文为基础.
15. Paulo Mancosu 和 Ezio Vailati,"Torricelli's Infinitely Long Solid and Its Philosophical Reception in the Seventeenth Century",*Isis* 82(1991),50 - 70,p.54.
16. Mahoney,The Mathematical Career,p.220.
17. 同上,p.221.
18. 同上,p.230.
19. 关于寻找幂和公式的努力的讨论见 Ivo Schneider,"Potenzsummenformeln im 17.Jahrhundert",*Historia Mathematica* 10(1983),286 - 296.
20. Brunschvicg 和 P.Boutroux 主编的 Blaise Pascal,Oeuvres(Paris:Hachette,1909),vol.Ⅲ,p.365.

21. Struik, Source Book, p.247. 关于约翰·沃利斯工作的更多材料见 J.F.Scott, *The Mathematical Work of John Wallis*(New York: Chelsea, 1981).
22. Evelyn Walker, *A Study of the Traité des Indivisibles of Gilles Persone de Roberval*(New York: Columbia University, 1932), p.174.
23. Struik, Source Book, p.239.
24. H.W.Turnbull 主编的 *James Gregory Tercentenary Memorial Volume*(London: G. Bell and Sons, 1939), p.148.
25. 同上, p.170.
26. R.C.Gupta, "The Madhava-Gregory Series", *The Mathematics Educator*(*India*)7(1973), 67 – 70, p.68. 20世纪关于印度幂级数最早的报告是 K.Mukunda Marur 和 C.T.Rajagopal 的"On the Hindu Quadrature of the Circle", *Journal of the Bombay Branch of the Royal Asiatic Society* 20(1944), 65 – 82 以及 C.T.Rajagopal 和 A.Venkataraman 的"The Sine and Cosine Power Series in Hindu Mathematics", *Journal of the Royal Asiatic Society of Bengal*, Science 15(1949), 1 – 13. 对反正切级数的证明的评论见 C.T.Rajagopal 的"A Neglected Chapter of Hindu Mathematics", *Scripta Mathematica* 15(1949), 201 – 209 以及 C.T.Rajagopal 和 T.V.Vedamurthi Aiyar 的"On the Hindu Proof of Gregory's Series", *Scripta Mathematica* 15(1951), 65 – 74. 对它作了讨论的还有 Ranjan Roy 的"The Discovery of the Series Formula for π by Leibniz, Gregory and Nilakantha", *Mathematics Magazine* 63(1990), 291 – 306. 正弦函数和余弦函数的级数在 Victor J.Katz 的"Ideas of Calculas in Islam and India", Mathematics Magazine 68(1995), 163 – 174 中有详细的讨论. 进一步的信息见 C.T.Rajagopal 和 M.S.Rangachari 的"On an Untapped Source of Medieval Keralese Mathematics", *Archive for History of Exact Sciences* 18(1978), 89 – 101 和"On Medieval Kerala Mathematics", *Archive for History of Exact Sciences* 35(1986), 91 – 99.
27. 对范·休莱特的最好论述见 J.A.van Maanen 的"Hendrick van Heuraet(1634 – 1660?): His Life and Mathematical Work", *Centaurus* 27(1984), 218 – 279.
28. J.M.Child, *The Geometrical Lecture of Issac Barrow*(Chicago: Open Court, 1916), p.117. 通过把巴罗的几何工作翻译成现代的分析术语 Child 在他的评论中似乎把发明微积分的主要功劳归给了巴罗,但巴罗自己从未迈出脱离开几何这一步. 不论如何,巴罗的讲义是很有价值的. 对该书的一个评论见 Florian Cajori 的"Who Was the First Inventor of the Calculus", *American Mathematical Monthly* 26(1919), 15 – 20.
29. 同上, p.39.
30. 同上, p.120.
31. Richard Westfall, *Never at Rest*(Cambridge: Cambridge University Press, 1980), p.x. 这部传记包含有激动人心的细节,它不仅论述了牛顿在数学上的成就,还包含了牛顿在科学其它各领域的工作. 对牛顿数学成就的一个绝佳综述见 V.Frederick Rickey 的"Isaac Newton: Man, Myth, and Mathematics", *College Mathematics Journal* 18(1987), 362 – 389. 对牛顿早期工作的一个综述见 Derek T.Whiteside 的两篇文章"Isaac Newton: Birth of a Mathematician", *Notes and Records of the Royal Society* 19(1964), 53 – 62 和"Newton's Marvelous Years: 1666 and All That", *Notes and Records of the Royal Society* 21(1966), 32 – 41.
32. John Fauvel, Raymond Flood, Michael Shorthand 以及 Robin Wilson 编辑的"Let Newton Be! *A new perspective on his life and works*(Oxford: Oxford University Press, 1988), p.15.
33. Westfall, *Never at Rest*, p.191.
34. 同上, pp.192 和 209.
35. Derek T.Whiteside, *The Mathematical Papers of Isaac Newton*(Cambridge: Cambridge University Press, 1967 – 1981), vol. Ⅲ, p.33 – 35. 所有现存的牛顿的数学手稿都被收集起来,适当时加以翻译,然后被编辑到这部八卷本的套书中. 这些集子值得仔细浏览. 对牛顿微积分的一个简要介绍见 Philip Kitcher, "Fluxions, Limits and Infinite Littleness: A Study of Newton's Presentation of the Calculus", Isis 64(1973), 33 – 49.
36. Turnbull, *Correspondence*, vol. Ⅱ, p.131. 牛顿将包含有对二项式定理这一讨论的信——后信——寄给了亨利·奥尔登堡,后者又将其交给了莱布尼茨. 它同 1676 年 6 月 13 日的前信一起成为后来皇家学会指控莱布尼茨剽窃的报告中所引用的核心证据.

37. Whiteside, *Mathematical Papers*, Ⅱ, pp. 241 – 243.

38. 同上, Ⅱ, p. 213.

39. 同上, Ⅲ, p. 71.

40. 同上, Ⅲ, p. 73.

41. 同上, Ⅲ, p. 75.

42. 同上, Ⅲ, p. 81.

43. 同上.

44. 同上.

45. 同上, Ⅲ, p. 83.

46. 同上, Ⅲ, p. 117.

47. 同上, Ⅲ, p. 151. 更多有关曲率的材料见 Julian L. Coolidge, "The Unsatisfactory Story of Curvature", *American Mathematical Monthly* 59(1952), 375 – 379.

48. 同上, Ⅲ, p. 237.

49. Westfall, *Never at Rest*, p. 403.

50. De Gandt, Force and Geometry, p. 18.

51. 同上, p. 22. 同上, p. 39.

52. 同上, p. 31.

53. 同上, p. 39.

54. 有关牛顿对平方反比力意味着椭圆轨道这一结果的证明的讨论见 Bruce Pourciau, "Reading the Master: Newton and the Birth of Celestial Mechanics", *American Mathematical Monthly* 104(1997), 1 – 19.

55. Isaac Newton, *Principia*, A. Motte 译 F. Cajori 修订(Berkeley: University of California Press, 1966), pp. 38 – 39.

56. 关于莱布尼茨的新近的著作中最出色的包括有: Eric Aiton, *Leibniz*, *A Biography*(Bristo: Adam Hilger Ltd, 1985) 和 Joseph E. Hofmann, *Leibniz in Paris*, 1672—1676(Cambridge: Cambridge University Press, 1974). 前者是论述莱布尼茨整个科学生涯的一部一般著作, 后者则则详细论述莱布尼茨发明他的微积分版本的岁月.

57. Child, *Early Mathematical Manuscripts*, pp. 30 – 31. 这是"Historia et Origo" 的一部分.

58. 同上, p. 100.

59. 同上, p. 143.

60. Struik, Source Book, p. 275.

61. 同上, p. 279.

62. G. W. Leibniz, *Mathematische Schriften*, C. I. Gerhardt 编(Hildesheim: Georg Olms Verlag, 1971), vol. Ⅴ, 320 – 328.

63. Johann Bernoulli, *Opera Omnia*(Hildesheim: Georg Olms Verlag, 1968), vol. Ⅰ, 179 – 187, p. 183.

64. Child, Early *Mathematical Manuscripts*, p138.

65. G. W. Leibniz, *Mathematische Schriften*, vol. Ⅴ, 294 – 301. 英文版见 Struik, Source Book, p. 282.

66. 同上, vol. Ⅴ, 285 – 288

67. H. J. M. Bos, "Differentials, Higher – Order Differentials and the Derivatives in the Leibnizian Calculus", *Archive for History of Exact Sciences* 14(1974), p. 1 – 90, p. 56. 作为这段引文出处的原始论文可以在 Leibniz, *Mathematische Schriften*, vol. Ⅴ. p. 350. 查到. Bos 的论文是对微分一般思想的一个绝佳研究.

68. 同上.

69. 对争论的详细讨论见 A. R. Hall 的 *Philosophers at War: The Quarrel Between Newton and Leibniz*(Cambridge: Cambridge University Press, 1980).

70. Struik, Source Book, p. 313. Struik 给出了洛必达《分析学》部分章节的翻译. 有关洛必达工作的更多细节见 Carl Boyer, "The First Calculus Textbook", *Mathematics Teacher* 39(1946), 159 – 167.

71. 同上, p. 315 – 316. 又见 Dirk Struik"The Origin of l'Hospital's Rule", *Mathematics Teacher* 56(1963), 257 – 260.

72. Humphry Ditton, *An Institution of Fluxions*(London: Botham, 1706), p. 1.

微积分概览

1340—1425	马德哈瓦(Madhava)	幂级数
1445—1545	尼拉康达(Kerala Gargya Nīlakantha)	幂级数
1530—1610	加斯特德维(Jyesthadeva)	幂级数推导
1571—1630	开普勒(Johann Kepler)	最大值、面积、体积
1584—1667	圣文森特的格利高里(Gregory of St. Vincent)	双曲线下面积
1596—1650	笛卡儿(René Descartes)	法线
1598—1647	卡瓦列里(Bonaventura Cavalieri)	面积和体积
1601—1665	费马(Pierre de Fermat)	极值、切线、面积
1602—1675	罗伯华(Gilles Persone de Roberval)	切线、面积
1608—1647	托利拆里(Evangelista Torricelli)	面积和体积
1615—1677	奥尔登堡(Henry Oldenburg)	微积分的通信人
1616—1703	沃利斯(John Wallis)	面积
1618—1667	撒拉萨(Alfonso Antonio de Sarasa)	对数和面积
1620—1687	梅卡托(Nicolaus Mercator)	对数的幂级数
1622—1685	斯卢兹(René François de Sluse)	导数的算法
1623—1662	帕斯卡(Blaise Pascal)	面积
1625—1683	约翰·柯林斯(John Collins)	微积分的通信人
1628—1704	约翰·胡德(Johann Hudde)	导数的算法
1630—1677	巴罗(Isaac Barrow)	面积、切线、弧长
1634—1660	范·休莱特(Hendrick van Heuraet)	弧长
1638—1675	詹姆斯·格利高里(James Gregory)	面积、级数
1642—1727	牛顿(Isaac Newton)	级数、流数、基本定理、天体力学
1646—1716	莱布尼茨(Gottfried Wilhelm Leibniz)	微积分、基本定理
1661—1704	洛毕达(Guillaume François l'Hospital)	微分学教科书
1667—1748	约翰·伯努利(Johann Bernoulli)	指数的微积分
1675—1715	狄顿(Humphry Ditton)	微积分教科书
1678—1760	海耶斯(Charles Hayes)	微积分教科书

第13章 18世纪的分析学

> 作为一位知名的数学教授,约·伯努利向全世界思想敏锐的数学家们致以深深的敬意.由于确实知道,几乎没有任何事物能够比提出困难的同时又有用的问题更能极大地激发那些高贵的和敏锐的心灵进行增进知识的工作,……通过解决这些问题,他们在获得荣誉的同时也给后人树立起永恒的丰碑,因此,我应该期待从数学界得到我应得的那份谢意,……如果我会在当代一流的分析学家面前提出一些问题,……而在这些问题上他们可以检验自己的方法,可以施展自己的才华,并且倘若他们揭示了些什么,可以与我们联系以使每个人能够公开地从我们这里获得他们应得的赞颂.
>
> ——1697年1月,在荷兰的格罗林根发出的公开宣言[1]

1739年3月30日,列昂哈德·欧拉(Leonhard Euler)向圣彼得堡科学院递交了一篇论文,在其中首次出现了对三角函数的微积分的讨论.此前从未有过被含有字母和数字的公式表示出的如同代数函数一般的正弦、余弦函数的概念,这些字母和数字对于其他这样的公式的关系可以通过微积分的手法加以研究.仅在那时欧拉才认识到,正弦和余弦函数作为来自于振动理论的微分方程的解很自然地出现了.它们可以和其他类型的函数相结合.欧拉通过给其他的数学家的信件使得他的发现为人所知,最后在1748年在他的《导论》(Introductio)中详细发表了这个内容.

驱动18世纪的微积分学不断发展的动力是解决物理问题的需要,物理问题的数学表达一般都是用微分方程的形式.牛顿和莱布尼茨都在曲线的研究中求解过微分方程,但他们之后的数学家们逐渐地将重点从研究曲线和与之相关的几何变量转移到研究含有一个或多个变量以及某些常量的解析表达式,即一个或多个变量的函数.由经常从应用牛顿运动定律中产生的某些物理状态所决定的,在这些变量的微分和依赖于它们的变量之间的关系产生微分方程,微分方程的解明显地决定了所需要的函数.实际上,通过它们所满足的微分方程可以发现和分析新的函数类.

在18世纪的分析学发展中的主要的人物是在历史上最多产的数学家欧拉.这一章的大部分将专注于他在微分方程理论、变分学和多元微积分的工作,以及他的三本有影响的分析学教科书.但是,本章将从伯努利家族给欧洲数学家提出的一些很有挑战性的问题开始,这些问题的解答有助于

在数学中建立新思想，这些问题后来由欧拉和其他的数学家所发展. 因为有影响的思想和技巧也出现在英格兰的柯林·麦克劳林(Colin Maclaurin)和汤普逊·辛普森(Thomas Simpson)以及意大利的玛利亚·盖塔娜·艾格纳丝(Maria Gaetana Agnesi)的著作中，这些著作也将被讨论到. 尤其是，麦克劳林的著作在一定程度上回答了乔治·伯克利(George Berkeley)对微积分学基础的批评，他对这种批评的反应也将讨论到. 本章将以约瑟夫·路易斯·拉格朗日(Joseph-Louis Lagrange)试图消除所有对无穷小量甚至极限的引用而把微积分学建立在幂级数的概念上作为结束.

13.1　微分方程

雅各布(Jakob)和约翰(Johann)伯努利兄弟，通常也称作雅克(Jacques)和让(Jean)或詹姆斯(James)和约翰(John)，处于欧洲理解莱布尼茨的新技巧并将之用到解决新问题的第一批人士之中. 例如，1659 年惠更斯(Huygens)已经用无穷小量发现并用几何证明，若一个物体受重力作用沿着曲线下降，从曲线上的任一点出发到达底部的时间都相等，则这条曲线是摆线. 惠更斯(Huygens)后来把这一思想用于摆钟的发明. 他认识到若摆被约束在摆线弧上运动，就会保持良好的准时性，无论摆动的尺度有多大. 雅各布·伯努利在 1690 年通过对于这一相等时间的曲线即等时线，建立微分方程，解析地证明了惠更斯的结果.

继成功地解决了等时线问题后，雅各布(图 13.1)提出了确定悬链线的形状的新问题，悬链线被假定为一条自由地悬挂在两个定点之间的柔软的无弹性的细绳. 伽利略曾经认为这条曲线是抛物线. 雅各布自己却无法解答这个问题. 但是在《教师学报》(Acta Eruditorun)的 1691 年 6 月的那一期上刊登了莱布尼茨、惠更斯和约翰·伯努利的解法. 约翰对于他能超过他的兄长感到非常自豪，据说解答这个问题时他彻夜未眠. 他后来在写给洛必达的讲义中更详尽地记述了他的解法. 他从使细绳保持位置的力的分析所得出的微分方程 $dy/dx = s/a$ 开始，这里 s 代表弧长. 又因为 $ds^2 = dx^2 + dy^2$，所以对原方程进行平方，得



图 13.1　瑞士邮票上的雅各布·伯努利.

$$ds^2 = \frac{s^2dy^2 + a^2dy^2}{s^2} \quad 或 \quad ds = \frac{\sqrt{s^2 + a^2}dy}{s},$$

或者，最后，

$$dy = \frac{sds}{\sqrt{s^2 + a^2}}.$$

积分后得到 $y = \sqrt{s^2 + a^2}$ 或者 $s = \sqrt{y^2 - a^2}$. 伯努利得出结论：

$$dx = \frac{ady}{s} = \frac{ady}{\sqrt{y^2 - a^2}}.$$

他虽然不能用闭形式来表达这个积分，但他却可以用某些圆锥截线来构造所求的曲线. 用现代术语来说，这个方程的解是 $x = a\ln\left(y + \sqrt{y^2 - a^2}\right)$ 或者 $y = a\cosh\frac{x}{a}$ 的形式，当然，在 1691 年，对伯努利和他的同时代的人来说，用已知曲线之下的面积或它的长度来作出回答已经足够了.

接下来的几年里，兄弟俩人提出其他问题，包括微分方程，并且同莱布尼茨一起在发展求解方法上取得了很大的进步. 尤其是 1691 年莱布尼茨发现了分离变量的技巧，亦即，一个微分方程可以改写成 $f(x)dx = g(y)dy$ 的形式并且对其两边进行积分后就得出解. 他也发展了求解齐次方程

$dy = f(y/x)dx$ 的用 $y = vx$ 进行替换然后分离变量的技巧.到1694年他还解决了一般的一阶线性微分方程 $mdx + nydx + dy = 0$,这里 m 和 n 都是 x 的函数.(用现代的记法,这是方程 $dy/dx + ny = -m$.)他通过 $dp/p = ndx$ 定义 p,替换得到 $pmdx + ydp + pdy = 0$.因为左边的后两项等于 $d(py)$,积分给出 $\int pmdx + py = 0$.这个方程给出了用面积表示的解答,使莱布尼茨得到了想要的答案.

人物小传

雅各布(1654—1705)和约翰·伯努利(1667—1748)(Jakob and Johann Bernoulli)

雅各布·伯努利自学数学,在法国、荷兰和英国生活,在那里他熟悉各种各样的科学家和数学家的工作.当他在1683年回到自己的出生地巴塞尔时,他学习了范·舒滕(Van Schooten)版的笛卡儿的几何并且最终被委任到巴塞尔大学,先在1687年是实验物理学的讲师,后是数学教授直到生命的尽头.与此同时,他的弟弟约翰试图从商以取悦父亲未果,在该大学学医,再次迎合父亲的愿望.然而,他却花很多时间同雅各布研究数学.他们一起掌握了莱布尼茨的早期工作,并且很快地能够作出自己的贡献.

约翰·伯努利在1690年代早期更详尽地在不同的文章中发展了莱布尼茨的技巧,因此使他能够通过惠更斯的帮助被任命为荷兰格罗林根大学数学系主席一职,直到1705年他哥哥去世使他继任巴塞尔数学系主席.

13.1.1 最速降线问题

或许就其对数学产生的最终结果而言,约翰·伯努利所提出的最有意义的问题是关于最速降线,即**最快下降的曲线的问题**.在《教师学报》的1696年6月期上他首次提出该问题如下,"请数学家解答的新问题:若 A 和 B 是垂直平面内给定的两点,确定运动粒子 M 的路径 AMB,沿着该路径 M 在自身重力的作用下下降,在最短的时间内从点 A 走到点 B."[2] 他注明这条曲线不是直线,而是一条"几何学家熟悉的"曲线.他要求答案不得晚于1696年底,但是1697年一月上旬,他听从莱布尼茨的劝告,将期限推迟到复活节,并正如在公开征答中所提及的将这个问题送给那些没有在《教师学报》上看到短函的人们.他提供一项奖励"并非金银,……,然而,由于美德本身就是它自己最渴望的回报并且名誉是一种有力的刺激,我们提供这种奖励 …… 复合着荣誉、赞赏和认可.因此我们将公开地和私下地、文字地和口头地加冕、尊敬和颂扬我们伟大的太阳神的敏锐."[3]在被送给挑战书的数学家中有牛顿,他被伯努利认为是窃取了莱布尼茨的方法,并且没有能力自己解答这个问题.

当牛顿在1月29日下午约4点接到伯努利的信时,他在造币厂劳累一天以后已是非常疲惫.尽管如此,他还是坚持到翌日清晨4点解决了这个问题后才去休息.伯努利不得不承认牛顿的才能.莱布尼茨完全为此事感到为难,他写信给皇家科学院否认他被牵扯到这件事中.不管如何,牛顿的解法与莱布尼茨、雅各布·伯努利和约翰的解法一起在《学报》的1697年5月期上刊出.

约翰·伯努利在解答这个问题时用到了两条物理定律.首先,他根据伽利略观点注意到,下落的物体的速度和下降的距离的平方根成正比.其次,他引用了史纳尔定律,当一束光线从稀薄的介质进入到稠密的介质时,光线会弯曲使得入射角正弦和折射角正弦的比值反比于两种介质的密度,因此正比于在两种介质中的速度.这个定律是费马由应用光经过的路径一定是费时最少的路径的原理所得出的.伯努利然后假设问题中的垂直平面是由密度不同的无穷小厚度的层构成.最速降线,费时最少的路径,因此就是光线的被弯曲的路径,其中光线的方向随光线从一层流动到下一层不断改变.在每一点上曲线的切线和纵轴的夹角的正弦与速度成正比,而速度又与下落的距离的平

方根成正比.

现在用 AMB 来表示所求的最速降线曲线,并用 AHE 代表在每一点速度的曲线,伯努利设 x 和 y 分别为从原点 A 到点 M 的纵坐标和横坐标,t 代表相应的速度点 H 的横坐标(图 13.2).当点 m 无限地接近点 M 时,他分别使用 dx、ds 和 dy 来表示无穷小量 Cc、Mm 和 nm.从折射角 nMm 的正弦等于 $dy:ds$,又正比于速度 t,伯努利得出方程 $dy:t=ds:a$ 或者 $ady=tds$,或者 $a^2dy^2=t^2ds^2=t^2dx^2+t^2dy^2$,或者,最后

$$dy=\frac{tdx}{\sqrt{a^2-t^2}}.$$

图 13.2 约翰 · 伯努利的最速降线问题.

因为曲线 AHE 是抛物线,有方程 $t^2=ax$ 或者 $t=\sqrt{ax}$,所以对 t 进行替换后就产生了曲线 AMB 的微分方程:$dy=dx\sqrt{\frac{x}{a-x}}$.

伯努利立即发现这个方程确定一条摆线.为了解析地证明这一点,他记

$$dx\sqrt{\frac{x}{a-x}}=\frac{adx}{2\sqrt{ax-x^2}}-\frac{(a-2x)dx}{2\sqrt{ax-x^2}}.$$

给定 $y^2=ax-x^2$ 为圆 GLK 的方程,以及明白上式的右端的第一项是沿着这个圆的弧长的微分之后,则方程积分给出 $CM=\text{arc}LK-LO$.因为 $MO=CO-CM=CO-\text{arc}GL+LO$.再假设 CO 等于圆周长的一半,就可以推出 $MO=\text{arc}LK+LO$ 或者 $ML=\text{arc}LK$.这就立即可以断定 AMK 是一条摆线.伯努利对于这条曲线与等时线相同表示惊喜.他注意到只因速度和距离的平方根成正比而不是和任何其他的力量这才是正确的."我们可以推测",他接着说,"大自然希望它是这样.因为大自然习惯于总以最简单的方式行进,所以在这里她通过一条相同的曲线就完成了两种不同的服务."[4]

虽然约翰 · 伯努利的最速降线问题的解法很巧妙,但是他的哥哥的解法更容易一般化.雅各布这样分析,如果沿着整条曲线一个质点以最短的时间运动,那么曲线的任何无穷小段将有同样的性质.利用一种几何的论证,他能够对曲线导出微分方程 $ds=\sqrt{a}dy/\sqrt{x}$,它容易转化成约翰的方程.[5] 雅各布的方法提供了数学的新领域,即在其中寻求满足某种最大值或最小值性质的变分学的开端.

另一方面,约翰对这一问题的解法引发他对曲线族的性质的研究,它转而又引出在多元函数理论中某些根本性的新概念.假定可以构造一族摆线 $\{C_\alpha\}$,其中每一条都是从给定点 A 到点 B_α 的最速降线,约翰 · 伯努利提出并解决了求一个新的曲线族 $\{D_\beta\}$ 的问题,它称为**等时线**,即任意一条曲

线上的每一点都是粒子以给定的时间 t_β 从点 A 沿着不同的摆线 C_α 下降到达的位置(图 13.3). 按物理术语,若摆线代表光线,则等时线代表波前,在相同时刻从点 A 发出的不同的光波的同时位置. 伯努利认识到,从光学理论(惠更斯发展的) 可以预言,等时线会和最速降线垂直相交. 因此他的几何问题就是求与过给定顶点的摆线族 $\{C_\alpha\}$ 正交的曲线族 $\{D_\beta\}$. 约翰几乎毫不困难地就能构造出等时线,但是他却为解决求一族给定的超越曲线的正交轨道的一般问题向别人挑战.

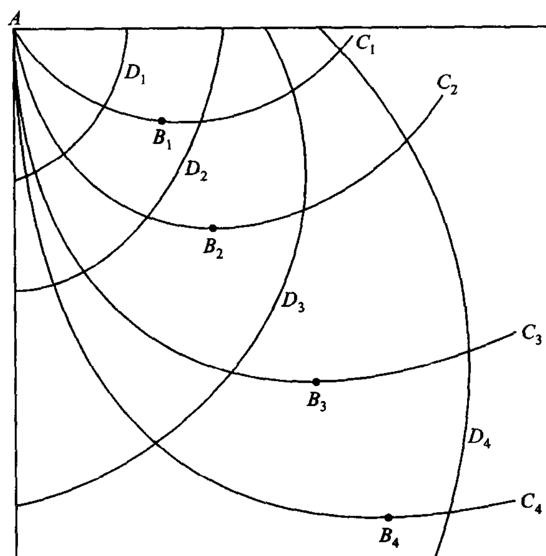


图 13.3 两个正交的曲线族.

13.1.2 二元函数的微分学

偏导数的最早观念是按照曲线族 $\{C_\alpha\}$ 的方式形成的,而不是像人们以现代的观点可能希望的那样,按照二元函数所定义的曲面的方式. 莱布尼茨在 1690 年代早期就已经开始考虑这样的曲线族了. 在基本情形中,在已知的曲线族里有两条无限接近的曲线 C_α 和 $C_{\alpha+\mathrm{d}\alpha}$,根据这个曲线族的几何定义的第三条曲线 D 与它们相交. 例如,D 可以垂直于曲线族中所有的曲线. 在这个情形下,要求出 D 的微分方程或者构造它的切线,就必需考虑坐标 y 的三个不同的微分. 设 P、P' 是 C_α 上的点及 Q、Q' 是 $C_{\alpha+\mathrm{d}\alpha}$ 上的点(图 13.4). y 的一个微分是在曲线族的单条曲线上的两点之间的微分,也就

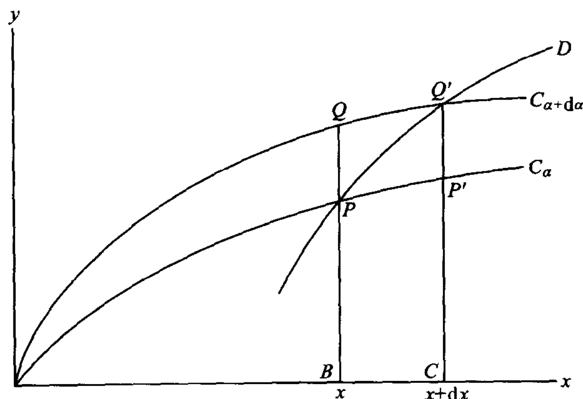


图 13.4 曲线族中的偏微分.

是 $y(P')-y(P)$. 这是当 x 为变量和 α 为常量时 y 的微分，记为 $d_x y$. 第二个微分是在相临的两条曲线上的对应点的 y 值之间的微分，即 $y(Q)-y(P)$. 这是当 α 为变量和 x 为常量时 y 的微分，记为 $d_\alpha y$. 这种微分的微分法被当作是曲线到曲线的微分. 最后，第三个微分 $y(Q')-y(P)$ 记为 dy，是沿着曲线 D 的微分.

对于莱布尼茨而言，若是曲线代数地给定，计算前两种微分都毫无困难. 他可以将 x 或 α 看作常量使用他的标准法则，实际上分别取对 x 或 α 的偏导数. 不幸的是，很多令人感兴趣的曲线不是代数地给定的，而是以积分的方式. 例如，上面的最速降线摆线表示为

$$C_\alpha = y(x,\alpha) = \int_0^x \sqrt{\frac{x}{\alpha - x}}\,dx.$$

(这里为了清楚起见，应用了关于积分符号的现代记号. 莱布尼茨和伯努利都没有用这样的记法.) 在一封被他带回家的 1697 年 7 月来自约翰 · 伯努利的信中，给莱布尼茨提出的问题是如何进行像这样一种情形的含有变量 α 的微分.

莱布尼茨经过几天的考虑后(与此同时打算在俄国彼得大帝从柏林到荷兰的途中会见他) 成功地解决了这个问题. 他的解法又回到了他自己的微积分的基本思想，即微分是有限差分的外推，而积分是有限和的外推. 因为对于有限数量的两个集合，部分的差分的和等于部分的和的差分，莱布尼茨发现了所谓的微分法和积分法的**可交换性定理**，即

$$d_\alpha \int_b^x f(x,\alpha)\,dx = \int_b^x d_\alpha f(x,\alpha)\,dx.$$

(这一定理的现代表述是以对于 α 的导数而不是微分的形式.) 莱布尼茨于是可以应用这个结论来求曲线 D 的切线，这里的曲线 D 在对数曲线族 $y(x,\alpha)=\alpha\log x$ 上截出相等的弧段.

因为对数曲线的弧长的微分 ds 是由 $ds = \frac{\sqrt{x^2+\alpha^2}\,dx}{x}$ 给出，要求的曲线 D 由条件

$$s(x,\alpha) = \int_1^x \frac{\sqrt{x^2+\alpha^2}}{x}\,dx = K$$

决定，这里 K 是常数. 为了确定切线，莱布尼茨需要计算在图 13.5 的微分三角形中的 QB 和 QB' 的比. 第一个值是直接的：$QB = d_\alpha \alpha \log x = \log x_0\, d\alpha$，这里 x_0 是 B 的 x 坐标. 由于 $AB = AB'$，弧 QB'

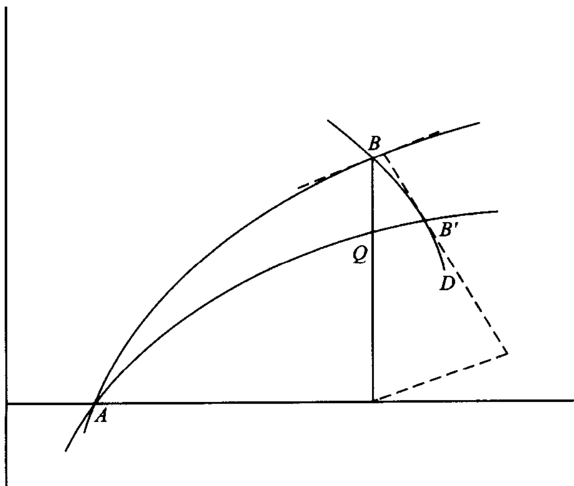


图 13.5　求在对数曲线 $y = 2\log x$ 上截出相等弧段的曲线 D 的切线.

等于差 $AB - AQ$.因此

$$QB' = d_\alpha s = d_\alpha \int_1^{x_0} \frac{\sqrt{x^2 + \alpha^2}}{x} dx.$$

由可交换性定理知,这个积分等于

$$\int_1^{x_0} d_\alpha \frac{\sqrt{x^2 + \alpha^2}}{x} dx = \int_1^{x_0} \frac{\alpha}{x\sqrt{x^2 + \alpha^2}} d\alpha dx.$$

因为积分的限只包含 x, 可是

$$QB' = \alpha d\alpha \int_1^{x_0} \frac{dx}{x\sqrt{x^2 + \alpha^2}}.$$

QB' 现在被表示成一特殊曲线下的面积,可以被当作已知;因此要求的比值也是如此.

莱布尼茨将他的解法送给伯努利,伯努利确信莱布尼茨的结果会使他在解决由含参数的积分所定义的曲线的正交轨道的问题上获得重要的进展.令人遗憾的是,尽管他找到了要求的曲线族的微分方程,但是对它积分的技术困难是不可克服的,伯努利始终未能发展利用莱布尼茨的结果的一种真正的一般方法.尽管如此,可交换性定理仍然成为二元函数微分学的基础之一.

二元函数微积分的另两个重要方面是由尼古拉斯·伯努利(1678—1759) 发现的一个全微分的概念和混合二阶导数的相等性.尼古拉斯·伯努利是约翰和雅各布两人的侄子,他生活在大名鼎鼎的叔伯的身影下,选择了很少发表数学文章.在 1719 年 6 月的《教师学报》上一篇文章中,尼古拉斯讨论正交轨道问题,但是他并没有给出证法并且直到几年后才以手稿的形式给出.在手稿中,他断言与一族含 α 的曲线族有关的微分 dy 是由 $dy = p dx + q d\alpha$ 给出,他称这个表达式为曲线族的**完全微分方程**,现在称为**全微分**.以现代术语来说,p 和 q 分别是 y 对 x 和 α 的偏导数.尽管伯努利没有说他如何推导这个结果,但一个可能的方法就是几何论证的方式.沿着与给定的曲线族相交的曲线 D 的微分 dy 可以被表为

$$\begin{aligned} dy &= y(Q') - y(P) = [y(Q') - y(Q)] + [y(Q) - y(P)] \\ &= d_x(y + d_\alpha y) + d_\alpha y = d_x y + d_x d_\alpha y + d_\alpha y. \end{aligned}$$

由于中间项,一个二阶微分,和其他项比较起来无限小,所以结论是 $dy = d_x y + d_\alpha y$ 或者 $dy = p dx + q d\alpha$(见图 13.4).我们在后边将看到,是欧拉最终使得微分系数 p 和 q 本身成为他的微积分的基础概念,取代了莱布尼茨和伯努利的微分.

在基于图 13.4 的另一个几何论证后得到了混合二次微分的相等性,在此图中一个为线段 $P'Q'$, 它可以通过两种不同的方法得到.一方面,$P'Q' = PQ + d_x PQ = d_\alpha y + d_x d_\alpha y$.另一方面,$P'Q' = d_\alpha CP'$,它又等于 $d_\alpha(BP + d_x BP) = d_\alpha BP + d_\alpha d_x BP = d_\alpha y + d_\alpha d_x y$. 比较两个表达式得到结果 $d_x d_\alpha y = d_\alpha d_x y$.非常有趣的是,尼古拉斯·伯努利并不认为这个论证是一个证明,而只是个对结果的图示说明而已."我想仅仅由微分观念来看,它对任何人都是显而易见的."[6] 他在自己的关于正交轨道的工作中运用这个成果,但是因为他的主要论证没有出现在他的出版著作中,它们的影响甚小.关于混合偏微分的相等性的定理于是大约在 10 年之后由欧拉重新证明.他注意到这个结果和可交换性定理,它们两者都是他和尼古拉斯·伯努利从相等性定理导出的,构成了偏导数理论的基础.对于由他的前辈们讨论过的求依据给定曲线族几何地定义的新曲线的同样的问题,欧拉运用它给出了他自己的解法.

欧拉在 1740 年前后发展的在微分方程解法上的另一个重要思想,也由 A.C 克莱罗(Alexis-Claude Clairaut,1713—1765) 在 1739 年独立发现,这就是对于齐次线性微分方程 $P dx + Q dy = 0$ 的可解性

的条件,其中 P 和 Q 是 x 和 y 的函数. 如果 $P\mathrm{d}x + Q\mathrm{d}y$ 是一个函数 $f(x,y)$ 的全微分,则由混合二阶微分的相等性得 $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$. 然而,更重要的是,克莱罗论证了这个条件对于 $P\mathrm{d}x + Q\mathrm{d}y$ 为全微分是充分的. 他断言,事实上,在这个条件下 $f(x,y)$ 由 $\int p\mathrm{d}x + r(y)$ 给出,其中 r 是 y 的待定函数. 从莱布尼茨的结论可知,$\int p\mathrm{d}x + r(y)$ 的微分等于 $p\mathrm{d}x + \mathrm{d}y\int \frac{\partial P}{\partial x}\mathrm{d}x + \mathrm{d}r$. 但是由于

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad 和 \quad \int \frac{\partial Q}{\partial x}\mathrm{d}x = Q + s(y),$$

微分又可以写成 $P\mathrm{d}x + Q\mathrm{d}y + \mathrm{d}r + s\mathrm{d}y$. 因此,如果取 r 满足 $\mathrm{d}r = -s\mathrm{d}y$,则微分变成了要求的 $P\mathrm{d}x + Q\mathrm{d}y$. 克莱罗就这样将原来的二元问题化简为一个他假定可解的一个变量的常微分方程. 他也容易把这个"恰当性"结果推广到多于二元的齐次线性方程.



图 13.6　瑞士邮票上的欧拉.

人物小传

列昂哈德 · 欧拉(1707—1783)(Leonhard Euler)

欧拉(图 13.6)生于瑞士的巴塞尔,很早就显露出他的卓越能力,15 岁就带着荣誉从巴塞尔大学毕业. 虽然他父亲更喜欢他为当牧师作准备,欧拉却设法使约翰 · 伯努利承诺私下指导他数学. 后者很快发现他的学生的天才,并说服欧拉的父亲允许他钻研数学. 在 1726 年欧拉被拒绝得到大学的一个位置,部分因为他年轻. 然而,几年以前,彼得大帝,鉴于莱布尼茨的促进,作为使俄国现代化的部分努力已经决定创建圣彼得堡科学院. 在 1725 年任命的科学院的最早的成员中有尼古拉斯 Ⅱ 世(1695—1726)和丹尼尔 · 伯努利(1700—1782),他们是约翰的两个儿子. 欧拉与两位建立了友谊. 尽管 1726 年在圣彼得堡得不到在数学方面的职位,他们却推荐他到医药和物理学的空位上. 此职位被欧拉立即接受了. (他在巴塞尔期间已经研究了这些学科.)

在 1733 年,在尼古拉斯去世和丹尼尔回到瑞士之后,欧拉被任命为科学院的首席数学家. 同年末他与凯瑟琳 · 吉塞尔结婚,同她接着有了 13 个孩子. 在那时的俄国一位外国科学家的生活并不总是无忧无虑的. 尽管如此,欧拉大体上能够机敏地脱离矛盾,直到 1741 年围绕俄国王位继承的问题使他确定接受普鲁士的弗里得里克 Ⅱ 世的邀请加入(也是在莱布尼茨建议下)弗里得里克 Ⅰ 世建立的柏林科学院. 他很快成为科学院数学部的指导者,随着他的分析学的课本以及大量的数学文章的发表,成为公认的欧洲首席数学家. 在 1755 年巴黎科学院授予他外籍院士,部分是表彰他 12 次赢得他们的两年一度的大奖赛. 然而,最终弗里得里克对欧拉缺乏哲学教养而感到厌倦. 当两人不能在资金安排和科学自由上取得一致时,欧拉在 1766 年受叶卡琳娜女皇之邀回到俄国. 她的继位标志着俄国回到彼得大帝的西化政策. 在对他的家庭有了经济保障下,欧拉继续他的数学活动,尽管他在 1771 年变得几乎完全失明了. 他的惊人的记忆使他能够在大脑中进行详细的计算. 因此他实际上能够口述文章和信件给他的儿子和其他人,直到在 1783 年的一天当他同他的一个孙子玩耍时突然去世.

13.1.3　微分方程和三角函数

尽管关于多元函数的思想被证明在解微分方程中很重要,欧拉在 1730 年代也发展了其他技巧,导致了正弦和余弦函数的现代符号的发明. 回忆在 1693 年莱布尼茨通过一个几何论证已经得

到了关于正弦函数的微分方程,并用待定系数法解方程得到了正弦函数的幂级数表达式.在18世纪早年,产生这种方程的物理问题一般通常是被几何地求解.比如方程

$$dt = \frac{c\,ds}{\sqrt{c^2 - s^2}},$$

通过几何位置的考虑可以对 t 解得 $t = c\arcsin\frac{s}{c}$,而不是 $s = c\sin\frac{t}{c}$. 导致产生这种类型的方程的问题之一,是约翰·伯努利在1728年考虑过的,就是如何确定一个物体在受到一个与它到给定点距离成正比的力的作用下的运动.可是,正弦函数在那时仍然被看作是一个给定半径的圆上的一条线而不是在现代术语中的函数,并且不是这些答案的一部分.

与"遗漏"的正弦函数相反,指数函数在1730年已经被欧拉熟悉了.实际上,在到达圣彼得堡之后不久,在一份他为用作讲义而写的关于微分计算的手稿中,欧拉注意到有两类函数,代数函数和超越函数.后一类仅由指数函数和对数函数组成,他着手讨论了它的性质.这样他知道 $y = e^{ax}$ 满足的微分方程是 $dy = ae^{ax}dx$, 或者用另一个方式来讲,$dy = ay\,dx$ 的解是 $y = e^{ax}$. 在1730年代早期的不同的论文中,欧拉运用指数函数的这一性质解答了其他微分方程.例如,他注意到方程

$$dz - 2z\,dv + \frac{z\,dv}{v} = \frac{dv}{v}$$

可以通过乘以"积分因子"$e^{-2v}v$ 得到 $e^{-2v}v\,dz - 2e^{-2v}zv\,dv + e^{-2v}z\,dv = e^{-2v}dv$. 因为左端是 vze^{-2v} 的微分,方程的解是

$$e^{-2v}vz = C - \frac{1}{2}e^{-2v} \quad 或 \quad 2vz + 1 = Ce^{2v}.$$

高阶方程也可以利用指数函数来求解,但是到1730年代中期欧拉认识到这些函数不是足够的.例如,1735年丹尼尔·伯努利写信给欧拉讨论一个关于弹性棒的振动的问题.这个问题产生了四阶方程 $k^4(d^4y/dx^4) = y$. 伯努利和欧拉都认识到 $e^{x/k}$ 是一个答案,但是伯努利写到这个解是"对于目前的情形这不够一般."[7] 欧拉能够使用幂级数方法求解此方程,但是却没有认出在产生的解答中的正弦或余弦.直到1739年欧拉才认识到正弦函数将使所给的这样的高阶方程能够得到封闭形式的解.那年3月30日欧拉向圣彼得堡科学院递交了一篇论文,在其中他解答了正弦驱动的谐波振子的运动的微分方程,也就是说,一个物体受两个力的作用,一个正比于距离;一个随着时间正弦地变化.问题的这个提法或许正是正弦作为一个时间的函数的最早应用,并且产生的微分方程

$$2a\,d^2s + \frac{s\,dt^2}{b} + \frac{a\,dt^2}{g}\sin\frac{t}{a} = 0$$

(其中 s 表示位置,t 表示时间)是最早应用此函数的这样的方程.欧拉的解答在两方面是有价值的.第一,作为一个特殊的情形,他删掉了正弦项并解方程 $2a\,d^2s + \frac{s\,dt^2}{b} = 0$, 或者乘以 $b\,ds$ 后解方程 $2ab\,ds\,d^2s = -s\,ds\,dt^2$. 对 s 积分得到 $2ab\,ds^2 = (C^2 - s^2)dt^2$,或者 $dt = \frac{\pm\sqrt{2ab}\,ds}{\sqrt{C^2 - s^2}}$,这是对反正弦的微分方程(取正号)或者反余弦的微分方程(取正号).由于欧拉感兴趣的是运动而不是时间,所以解反余弦的方程求 s 而不是 t,得 $s = C\cos(t/\sqrt{2ab})$,这是历史上的第一个这样的显式的解析解.第二,为了解决一般的情形,欧拉假定有一个形如 $s = u\cos\frac{t}{\sqrt{2ab}}$ 的解,其中 u 为一个新变量.他再将这个解代入原方程求出 u. 这种处理表现出欧拉已经熟悉对于正弦和余弦的基本微分法则了.这些实际上莱布尼茨已经知道,如下边将提到的一样,它们已经出现在若干印刷资料中.

还有很多关于正弦函数和余弦函数的轶事.1739年5月5日欧拉给约翰·伯努利写到,他已经

用有限形式求解了三阶微分方程 $a^3 d^3 y = y dx^3$:"尽管看起来它很难积分,需要三次积分并且需要圆和双曲线的求积,但是它可以简化为一个有限的方程,积分的方程为

$$y = b e^{x/a} + c e^{-x/2a} \sin \frac{(f+x)\sqrt{3}}{2a}$$

…,这里 b, c, f 代表由三次积分产生的任意常数."[8]欧拉没有展现他是如何发现这个解的,但是一种适当的猜测是他用已知的指数形式的解 $y = e^{x/a}$ 来降低方程的阶数.在这个欧拉早先已经用过的技巧中,用 $e^{-x/a}$ 乘以原方程 $a^3 d^3 y - y dx^3 = 0$ 并假设它是 $e^{-x/a}(A d^2 y + B dy dx + C y dx^2)$ 的微分.于是直接证明原方程的新解也满足二阶方程 $a^2 d^2 y + a dy dx + y dx^2 = 0$.为了求解后面的这个方程需要用到一个不同的欧拉技巧.亦即,猜测一个形如 $y = u e^{\alpha x}$ 的解并用之代替方程中的 y.经过一点计算又表明令 $\alpha = -(1/2)a$ 可以消掉 $du dx$ 项.方程然后化简为 $a^2 d^2 u + (3/4) u dx^2 = 0$,它与欧拉在三月求解的那个方程具有相同的形式.在这一情形,解为 $u = C\sin((x+f)\sqrt{3}/2a)$,由此得到原三阶方程的通解.注意这种重构用到双曲线(指数函数,自然地与定义为双曲线下方的面积的对数函数有关)的求积分和圆(正弦函数,与定义中含有圆的面积的反正弦函数有关)的求积分,以及三个积分.

由于正弦函数和指数函数已经被用于相同的微分方程的求解,这清楚地表明欧拉现在认为正弦函数,以及通过扩展的其他三角函数,是像指数函数一样的在同样的意义下的函数.但是更有趣地,正是这些函数在微积分中的引入,将欧拉引向对于常系数线性微分方程的求解方法,即这种形式的方程:

$$y + a_1 \frac{dy}{dx} + a_2 \frac{d^2 y}{dx^2} + a_3 \frac{d^3 y}{dx^3} + \cdots + a_n \frac{d^n y}{dx^n} = 0.$$

欧拉 1739 年 9 月 15 日在给约翰·伯努利的信中写道,"通过用多种方法处理这个问题后,我完全意想不到地偶然发现了我的这个解法;在此之前我毫不怀疑代数方程的解在这个问题上有如此之高的重要性."[9]欧拉所说的"意外"的解法就是用代数方程 $1 + a_1 p + a_2 p^2 + \cdots + a_n p^n = 0$ 来代替所给的方程,因式分解"特征多项式"成为不可约的实的一次和二次因式之积.对于每个一次因式 $1 - \alpha p$ 取解 $y = A e^{x/\alpha}$,对于每个不可约的二次因式 $1 + \alpha p + \beta p^2$ 取解

$$e^{-\alpha x/2\beta}\left(C\sin \frac{x\sqrt{4\beta - \alpha^2}}{2\beta} + D\cos \frac{x\sqrt{4\beta - \alpha^2}}{2\beta} \right).$$

通解就是对应于每个因式的解的总和.作为一个例子,欧拉求解了丹尼尔·伯努利约四年前提出的一个方程

$$y - k^4 \frac{d^4 y}{dx^4} = 0.$$

对应的代数方程 $1 - k^4 p^4$ 因式分解为 $(1 - kp)(1 + kp)(1 + k^2 p^2)$,因此它的解就是

$$y = A e^{-x/k} + B e^{x/k} + C\sin \frac{x}{k} + D\cos \frac{x}{k}.$$

欧拉没有说他怎样得到他的代数求解方法.但是由于他在几个月前已经发现了三角函数被包含于方程 $y - a^3 \dfrac{d^3 y}{dx^3} = 0$ 的解中,可以猜测他只是将这个方法一般化.对于方程的给定的一个形如 $y = e^{x/a}$ 的解,约化过程表明本质上提供了特征多项式的一个因式分解 $1 - a^3 p^3 = (1 - ap)(1 + a^2 p^2 + ap)$.于是会容易得出欧拉九月的信里表明的一般的因式分解法.尤其会清楚正弦和余弦项来自于不可约的二次因式.

约翰·伯努利对欧拉的解法感到有些困惑.他注意到特征方程的不可约的二次因式可以在复

数域上因式分解，因而欧拉的解法可能产生与包含正弦和余弦函数的实解有关的特征多项式的复根．欧拉终于使伯努利相信 $2\cos x$ 和 $e^{ix}+e^{-ix}$ 是相等的，因为它们满足相同的微分方程，因此应用虚指数函数等同于应用正弦和余弦函数．同时也得到复指数函数与正弦和余弦函数有如下关系：

$$e^{ix}=\cos x+i\sin x \quad 和 \quad e^{-ix}=\cos x-i\sin x.$$

e^{-ix} 和 e^{ix} 的公式也有助于澄清在负数的对数的情形上的一个矛盾．如果这样定义的话，就会出现，例如，$2\log(-1)=\log(-1)^2=\log 1$，或者更一般地，$\log(-x)^2=\log(+x)^2$，由此 $2\log(-x)=2\log(+x)$，并且 $\log(-x)=\log(+x)$．实际上，准确地说约翰·伯努利辩论的是这个问题，即对数曲线一定是关于竖轴对称的，是一个似乎由等式

$$\frac{d(\log(-x))}{dx}=\frac{1}{x}=\frac{d(\log(+x))}{dx}$$

所确定的位置．但是欧拉公式表明 $ix=\log(\cos x+i\sin x)$，并且因此一个给定的(复)数有无穷多的对数．例如，$\log 1=\log(\cos 2n\pi)=2ni$，这里 n 可以是任意整数．因为，类似地，$\log(-1)=\log(\cos(2n+1)\pi)=(2n+1)\pi i$，在从一个集合而来的对数的两倍等于从另一个集合而来的对数的意义下，$2\log(-1)=\log 1$ 实际上是正确的．于是，正如欧拉所写的那样，他的方法使得处理负数和复数的对数的“所有困难和矛盾消失了．”[10]

13.1.4 变分学

欧拉对于分析学的有关微分方程的另一个领域即变分学做出了重大的贡献．这门学科是在目的为寻找使得某个特定积分取得最大值或最小值的曲线的问题的思考中成长起来的．例如最速降线问题就是寻找一条曲线使得在重力作用下下降的时间最短，也就是一条曲线，使得

$$I=\int dt=\int\frac{ds}{v}=\int\frac{\sqrt{1+y'^2}\,dx}{\sqrt{2gy}}$$

最小，这里 dt 是时间元素，ds 是弧长元素，g 是重力加速度，$y=y(x)$ 表示曲线，$v=\sqrt{2gy}$ 表示物体在重力影响下的下降速度．

在研究了这个问题和许多相近的问题之后，欧拉可以把它们都合并成一个普遍的原理，它确定一条使得一个形如 $I(y)=\int_a^b F(x,y,y')dx$ 的积分取得最大值或最小值的曲线 y．这个原理出现在他的较早期的数篇论文中，后来出现在他 1744 年的经典著作《求具有某种最大值和最小值性质的曲线的方法》中．欧拉的中心思想是使用一个对积分的折线逼近去发展一个取极值的必要条件．我们用现代术语给他的方法作一个简单的概括．

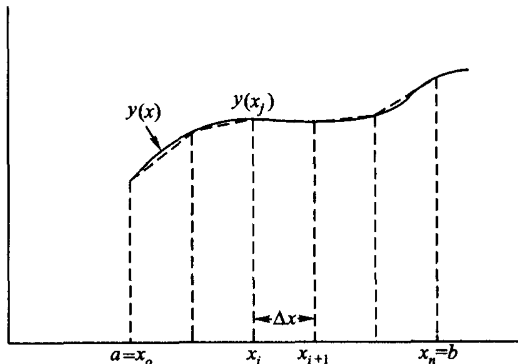


图 13.7 欧拉对 $y=y(x)$ 的多角形逼近．

将 $[a,b]$ 分成 n 个相等的子区间 $[x_{i-1},x_i]$，其中 $x_i=x_{i-1}+\Delta x$，$i=1,\cdots,n$，并考虑连接点 (x_i,y_i) 的多角曲线，这里 $y_i=f(x_i)$(图 13.7)．于是 $I(y)$ 可以被

$$I(y) \approx I(y_1, y_2, \cdots, y_{n-1}) = \sum_{i=0}^{n-1} F\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right)\Delta x$$

逼近.因为曲线 y 是极值,所以 I 对于每个 y_i 的偏导数必定为零,就是说曲线的"角点"的任何变化都一定导致积分失去它的极值性质.因此对每个 i 从 1 到 $n-1$,都有方程

$$\begin{aligned}
0 &= \frac{1}{\Delta x}\frac{\partial I}{\partial y_i} \\
&= \frac{\partial F}{\partial y}\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right) + \frac{\partial F}{\partial y'}\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right)\left(\frac{-1}{\Delta x}\right) + \frac{\partial F}{\partial y'}\left(x_{i-1}, y_{i-1}, \frac{y_i - y_{i-1}}{\Delta x}\right)\left(\frac{1}{\Delta x}\right) \\
&= \frac{\partial F}{\partial y}\left(x_i, y_i, \frac{y_{i+1} - y_i}{\Delta x}\right) - \left(\frac{1}{\Delta x}\right)\left[\frac{\partial F}{\partial y'}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) - \frac{\partial F}{\partial y'}\left(x_{i-1}, y_{i-1}, \frac{\Delta y_{i-1}}{\Delta x}\right)\right] \\
&= \frac{\partial F}{\partial y}\left(x_i, y_i, \frac{\Delta y_i}{\Delta x}\right) - \frac{\Delta\left(\frac{\partial F}{\partial y'}\right)}{\Delta x}.
\end{aligned}$$

当有无穷多个子区间和角点时,这些方程可以被代替为单个微分方程

$$\frac{\partial F}{\partial y} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = 0,$$

这是变分问题的基本必要条件,现在称为欧拉方程. 欧拉本人用微分符号来写他的方程:如果 F 是 x, y 的函数,并且 $p = \mathrm{d}y/\mathrm{d}x$,使得 $\mathrm{d}F = M\mathrm{d}x + N\mathrm{d}y + P\mathrm{d}p$,则积分当 $N\mathrm{d}y - p\mathrm{d}P = 0$ 时将取极值.但是,因为 $\mathrm{d}y = p\mathrm{d}x$,这可以被改写为 $N\mathrm{d}x - \mathrm{d}P = 0$ 或 $N - \mathrm{d}P/\mathrm{d}x = 0$.

我们给出两个例子来说明欧拉方程的应用方法. 首先,熟知两点之间的最短距离就是一条直线. 要通过变分的微分来得到它,我们必须对于两点之间的任意曲线使得 $\mathrm{d}s$ 最小,即我们必须找到使得 $I = \int \sqrt{1 + y'^2}\mathrm{d}x$ 最小的 y. 由于 $F(x, y, y') = \sqrt{1 + y'^2}$,因而有 $\partial F/\partial y = 0$,欧拉方程化为

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) = 0 \quad 或 \quad \left(\frac{\partial F}{\partial y'}\right) = c \quad 或 \quad \frac{y'}{\sqrt{1 + y'^2}} = c.$$

但是最后一个方程化为 $y' = a$,对某个常数 a,因此所求的曲线形如 $y = ax + b$,也就是一条直线.

我们的第二个例子又由欧拉给出,是最速降线问题. 在此情形下,由于函数 $F = \sqrt{1 + y'^2}/\sqrt{2gy}$ 只是 y 和 y' 的函数,修改欧拉方程去处理这种情况是最容易的. 先注意在此情形中

$$\frac{\mathrm{d}F}{\mathrm{d}x} = y'\frac{\partial F}{\partial y} + y''\frac{\partial F}{\partial y'}.$$

由欧拉方程,

$$\frac{\partial F}{\partial y} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right), \quad 所以 \quad \frac{\mathrm{d}F}{\mathrm{d}x} = y'\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F}{\partial y'}\right) + y''\frac{\partial F}{\partial y'} = \frac{\mathrm{d}}{\mathrm{d}x}\left(y'\frac{\partial F}{\partial y'}\right).$$

得到

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(F - y'\frac{\partial F}{\partial y'}\right) = 0 \quad 或 \quad F - y'\frac{\partial F}{\partial y'} = c.$$

取有兴趣的情形,后一个方程变成

$$\frac{\sqrt{1 + y'^2}}{\sqrt{2gy}} - \frac{y'^2}{\sqrt{2gy}\sqrt{1 + y'^2}} = c \quad 或 \quad \frac{1}{\sqrt{2gy}\sqrt{1 + y'^2}} = c \quad 或 \quad \frac{1}{2gy(1 + y'^2)} = c^2.$$

如果我们令 $a = \frac{1}{2gc^2}$,最后一个方程可改写为 $1 + y'^2 = \frac{a}{y}$ 或 $y'^2 = \frac{a - y}{y}$. 因此

$$\mathrm{d}y = \sqrt{\frac{a - y}{y}}\mathrm{d}x \quad 或 \quad \mathrm{d}x = \mathrm{d}y\sqrt{\frac{y}{a - y}},$$

将 x,y 互换后,就是约翰·伯努利对摆线所求出的相同的方程.[11]

欧拉最后认识到欧拉方程的推导并不是完全适当的,并且不能给出取极值的一个充分条件,但是1755年在他收到给出一个推导他的方程的更好的方法的来信之后,他赞扬了拉格朗日,将拉格朗日的方法递交给柏林科学院,留给年轻人去进一步探索这个领域.

13.2 微积分学课本

虽然有些微积分学课本是在17世纪和18世纪之交前后写的,包括那些在第12章中谈到的,但是在欧洲大陆和英国所看到的更多的是在该世纪的中间三分之一时期所写的.它们包括给非专业人员的用本国语言写的课本,以及为受过大学教育的人员而设计的用拉丁文写的一系列重要的课本.

13.2.1 托马斯·辛普森的《流数论》

在英格兰,中产阶级对数学知识的不断增长的需求被一群为辅助教学而写课本的私人教师部分地满足了.一个典型的例子就是托马斯·辛普森(Thomas Simpson)(1710—1761),他的最早的课本,《流数新论》,就是在他的私人学生的捐助下在1737年出版的.

辛普森的《流数论》在方法上基本上是牛顿的,大量地运用无穷级数去解决,尤其是积分中的问题.书中写满了问题,其中很多问题已经成为如今的学生所熟悉的.于是,在较前的关于最大值和最小值的一节中,辛普森展示了如何求内接于一个三角形的最大的平行四边形,外接于一个给定的圆的最小的等腰三角形,以及具有给定的体积的表面积最小的圆锥体.或许最早的确定一个多元函数的最大值的问题的解法也是包含在这一节中,此多元函数是 $w = (b^3 - x^3)(x^2z - z^3)(xy - y^2)$.虽然辛普森没有利用偏导数的语言,但是,在对每个情形令 $\dot{w} = 0$ 和同时求解所得的各个方程之前,他确实分别计算了 w 对 x, y 和 z 的数流关系,并保持其他的两个变量为常数.书的较后的一节致力于一个关于航行的问题,它或许是最早将正弦函数的微分规则刊出在课本中的.用辛普森的话来说,"任意圆弧的流数和他的正弦值的流数之比,等于半径和余弦值之比."[13] 这个证明,早在20多年前就由牛顿的《原理》的第二版的编辑,罗杰·柯泰斯(1682—1716),用相似三角形给出了.如果用 z 来表示一个半径为 An,圆心在 A 的圆上的一段圆弧,$x = Ab$ 表示 z 的正弦,bn 表示它的余弦,那么微分三角形 nrm,其中它的斜边 nr 表示 $\dot{z}$(圆弧的流数),它的边 mr 表示 $\dot{x}$(正弦的流数),相似于三角形 Anb(图13.8).这就得到被引用的结果 $\dot{z} : \dot{x} = An : bn$.用现代的术语来说,取 $An = 1$,这个式子可以写成

$$d(\sin z)/dz = \cos z.$$

辛普森因为由冠以他的名字的抛物线逼近所得的数值积分法则在现今最负盛名.这个法则不是出现在他的微积分课本中,而是在他的1743年的《关于种种物理和分析课题的数学论文》中.可是,它不是源于辛普森,而是已出现于甚至是17世纪的其他作者的著作中.

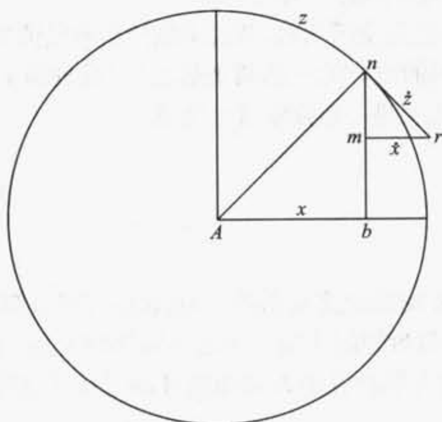


图 13.8 正弦的流数.

人物小传	托马斯 · 辛普森(1710—1761)(Thomas Simpson)
	辛普森出生于距离伯明翰不远的马凯特波斯华兹村.他的父亲要把他培养成一名织工.可是,托马斯对接受更好的教育的渴望导致他和他的父亲失和,他被逐出家门.到 25 岁他已经自学数学,包括来自洛必达著作的英译本的关于微积分的内容.在 1735 年他搬到伦敦,加入斯比塔 · 菲尔兹的数学会,一个现在位于居住郊区的纺织团体.这个社会团体的规则是使每位成员都有这样的责任"如果他被另一位成员问及数学或哲学问题,那么就用他所能够用的最粗浅的和最简单的方式去教他."[12] 通过他在团体中的活动,辛普森成为一名数学教师,并且很快能够放弃纺织并把家接到伦敦.他在数学界的声望通过几本教科书的出版而提高,最终使他能够在 1743 年在伍尔威奇皇家军事学院获得数学教授的职位.那是一所为了给军校学员成为工程师提供足够的数学教育而设立的学校.此后不久辛普森被选入皇家学会.

13.2.2 柯林 · 麦克劳林的的《流数论》

苏格兰数学家麦克劳林(1698—1746)的名字也为今天的学生从微积分课本中一个不是源于他的概念的麦克劳林级数所熟悉.这个级数出现在麦克劳林的《流数论》中,此书在 1742 年出版,部分原因是为了回击乔治 · 伯克利(1685—1753)在 8 年前道出的对于流数理论基础的批评.(这一批评的讨论见 13.5.1 节.)此著作第一册从几何的观点来处理牛顿的微积分的基础,但是在第二册中,麦克劳林有一个不同的安排,去展现流数法则以及它们以代数和算法方式的应用.麦克劳林于是给出了运用微积分的全部问题的细节.他讨论了极大值、极小值和拐点,发现了切线和渐近线,确定了曲率,并给出最速降线问题的一个完整的解释.麦克劳林计算了在由用 x 来表示的 y 所给定的曲线的下方的面积,方法是证明此面积的流数是 $y\dot{x}$,之后运用几种方法之一来确定这个表达式的流数.类似地,他通过先确定它们的流数计算了旋转体的体积和表面积.他运用了多重积分的一个基本形式来研究椭球的万有引力.最后,麦克劳林在处理对数时是从纳皮尔的用运动来表示的原始的定义开始的.那时,对他而言,确定对数的标准的流数的性质并运用这些性质来计算指数函数的流数是容易的.

人物小传	柯林·麦克劳林(1698—1746)(Colin Maclaurin)
	麦克劳林与辛普森不同,既受过大学教育又有大学事业.他出生于位于苏格兰西部的基尔摩丹村.他11岁进入格拉斯哥大学并很快掌握了大学数学课程.在19岁麦克劳林被任命为阿伯丁大学的数学系主任,但是此后不久动身赴三年的欧洲之旅去当一个富裕贵族之子的私人教师.阿伯丁当局对他的缺席并不特别高兴并且在他返回后不久迫使他辞职.在此期间,牛顿已经推荐他去爱丁堡的一个职位.他把一生的剩余的时光留在那里,教授的课程从欧几里得和初等代数跨越到流数和牛顿的《原理》.在1745年麦克劳林帮助保卫爱丁堡抵抗邦尼子查理的部队,但是当这座城市陷落时,他逃到约克.他在那里患病并再也没有康复,卒年48岁.

麦克劳林的工作比辛普森的工作和任何更早的工作含有更多的关于三角函数的某些内容.例如,除了关于正弦的辛普森定理以外,他几何地证明了圆弧的流数与其正切的流数之比等于半径和余弦之二重比,并且圆弧的流数与其正割的流数之比等于半径的平方与被正割和正切所确定的矩形之比.尽管这些结果可以被翻译成为现代的微积分定理($\frac{\mathrm{d}}{\mathrm{d}x}\tan x = \frac{1}{\cos^2 x}$;$\frac{\mathrm{d}}{\mathrm{d}x}\sec x = \sec x\tan x$),但是麦克劳林他们只给出与表示三角函数的那些线段有关的流数的比值.这些结果没有像关于对数函数和指数函数的微积分的那些结果一样被分析地应用.麦克劳林的结果包含反三角函数,但是它是被这样运用的.它们出现在求给定流数的流变(或积分)的内容中.因此麦克劳林注意到$\frac{\dot{y}}{a^2+y^2}$的流变是在半径为a的圆中正切为y的一段圆弧,而$\frac{a\dot{y}}{\sqrt{a^2-y^2}}$的流变是正弦为$y$的一段圆弧.而更有意思的是,麦克劳林也认识到在函数上的一个小的改变把来自一段圆弧的流变变成对数.例如,在第一个问题中在y^2的符号上的改变把流变变成$\frac{1}{2a}\log\left(\frac{a+y}{a-y}\right)$.因此似乎圆弧可以被虚对数所表示.相同的思想早已出现在柯泰斯的工作中了,但是柯泰斯和麦克劳林都不能得到由欧拉所搞出的完整的类推.

命名为麦克劳林的级数出现在第二册中:假设y可以表示成z的级数,记作$y = A + Bz + Cz^2 + Dz^3 + \cdots$.如果$E, \dot{E}, \ddot{E}, \cdots$是$y$的值和当$z$等于零时$y$的各阶流数,那么级数可以写成

$$y = E + \dot{E}z + \frac{\ddot{E}z^2}{1\times 2} + \frac{\dddot{E}z^3}{1\times 2\times 3} + \cdots$$

(假设$\dot{z}$等于1).麦克劳林的证明很简单,他假设y可以写成幂级数,即他先设$z = 0$,求出$A = E$.然后他求级数的流数,并再令$Z = 0$.得到$B = \dot{E}/\dot{z} = \dot{E}$.他继续求流数并令$z = 0$来完成结论.麦克劳林注解到,这个定理已经被布鲁克·泰勒发现并且在1745年发表在他的《增量的方法》(method of increments)中.

麦克劳林作出了很多这种级数的例子,包括在半径为a的圆中正弦和余弦函数的级数.例如,若$y = \cos z$(在半径为a的圆中),则$\frac{\dot{y}}{\dot{z}} = \sqrt{a^2-y^2}/a$,这就得到

$$\frac{\dot{y}}{\dot{z}} = \frac{a^2-y^2}{a^2},$$

和

$$\frac{2\dot{y}\ddot{y}}{\dot{z}^2} = -\frac{2y\dot{y}}{a^2} \quad 或 \quad \frac{\ddot{y}}{\dot{z}^2} = -\frac{y}{a^2}.$$

因此,由于当 $z = 0$ 时 $y = a$,我们得到 $E = a, \dot{E} = 0$ 和 $\ddot{E} = -\frac{1}{a}$.由于 $y = \cos z$ 级数的前三项就是 $y = a + 0z - \frac{1}{2a}z^2$.不必计算正弦和余弦的导数就容易求出更多的项.

为了确定极大值和极小值,麦克劳林也运用他的级数发展标准的导数判别:"当纵坐标的一阶流数为零,若同时它的二阶流数为正,则纵坐标为极小值;而若它的二阶流数为负,则纵坐标为极大值."[14] 若纵坐标 $AF = E$ 并且给定横坐标的两个值一个在 A 的右侧(记为 x),一个在左侧等距(记为 $-x$)(图 13.9),麦克劳林级数表明相应的纵坐标为

$$PM = E + \dot{E}x + \frac{\ddot{E}x^2}{2} + \frac{\dddot{E}x^3}{6} + \cdots,$$

和

$$pm = E - \dot{E}x + \frac{\ddot{E}x^2}{2} - \frac{\dddot{E}x^3}{6} + \cdots.$$

假设 $\dot{E}$ 且 x 足够小,麦克劳林得出结论,当 $\ddot{E}$ 为正(因此 AF 为极小值)时,这些纵坐标都将超过纵坐标 $AF = E$;并且当 $\ddot{E}$ 为负(因此 AF 为极大值)时,这些纵坐标都将小于纵坐标 AF.进而麦克劳林得出结论,若 $\ddot{E}$ 也为零而 $\dddot{E}$ 不为零,则或者 $PM > AF$ 且 $pm < AF$,或者反过来. 因此 AF 既非极大值也非极小值.

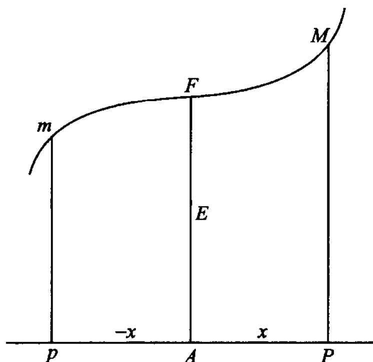


图 13.9 麦克劳林与第二导数判别.

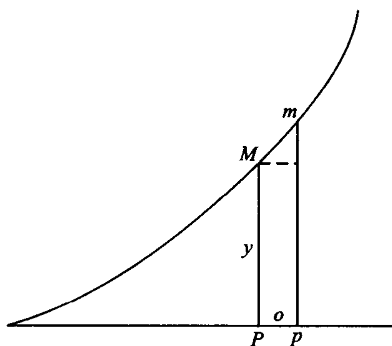


图 13.10 麦克劳林与微积分基本定理.

麦克劳林以部分的微积分基本定理的或许最早的解析证明结束他的课本,至少对于幂函数这种特殊情形是最早的.(他在课本中早已给出一个更一般的几何证明.)"假设 n 是任意(正)整数,… 如果在底 AP 或 x 上的面积总等于 x^n,则纵坐标 PM 或 y 应该总等于 nx^{n-1}."[15] 麦克劳林开始取底 x 的增量 $o = Pp$,并且注意到 $PM \times Pp = yo <$ 面积 $PMmp = (x + o)^n - x^n$(图 13.10).因为由他早先证明的代数结果 $(x + o)^n - x^n < n(x + o)^{n-1}o$,就得到 $yo < n(x + o)^{n-1}o$.类似地,用 P 左边的一个横坐标值,麦克劳林发现 $yo > n(x - o)^{n-1}o$,因此 $n(x - o)^{n-1} < y < n(x + o)^{n-1}$.不用现代的极限论述来证明 $y = nx^{n-1}$ 的麦克劳林使用了反证法论证.若 $y > nx^{n-1}$,则 $y = nx^{n-1} + r < n(x + o)^{n-1}$,对任意增量 o 成立.但若取 o 等于 $\left(x^{n-1} + \frac{r}{n}\right)^{1/(n-1)} - x$,经过简短的计算有 $y = n(x + o)^{n-1}$,矛盾. 若假设 y 小于 nx^{n-1},则出现相似的矛盾,麦克劳林的证明完成了.

13.2.3 玛利亚·艾格纳丝的分析学基础

人物小传	玛利亚·艾格纳丝(1718—1799)(Maria Agnesi)
	艾格纳丝是一位富有的米兰商人的最年长的孩子,她父亲通过聘请各种不同的教授个别指导她,鼓励女儿追求科学兴趣.到11岁她已经可以用几种语言流利表达,在十几岁她可以在诸如力学、逻辑学、动物学以及矿物学领域与当时最优秀的学者辩论重要的问题.在学过那时主要的数学著作之后,她开始教授她的弟弟们这门课程.她很快决定她的著作,包括代数的内容以及微分和积分学的完整的论文集,应该为了使所有意大利青年受益而出版.这个课本写得是那样清楚,使得法国科学院的一个委员会在1749年授权将它译成法文时写到"没有别的,用任何语言写的书籍,能够使得读者如此深入地,或如此迅速地进入分析学的基本概念."[16]约翰·柯尔森,18世纪中叶的剑桥的卢卡斯数学教授,对书的印象是如此深刻,以至于他学习意大利文的惟一目的是为了把这本著作译成英文,使得不列颠青年会像意大利青年一样受益. 鲍博也认识到艾格纳丝的才能并任命她为博洛尼亚大学的数学系主任,但是艾格纳丝从未接受这个职务.在1752年她父亲去世后不久,她从科学追求中退出并将她的余生用于宗教研究和在穷人中的社会工作.

麦克劳林的《流数论》在欧洲大陆上传阅,特别在1749年被译成法文之后.然而,在此前一年,第一个重要的洛必达的课本的继承者,在欧洲出现了,即艾格纳丝(1718—1799)的《给意大利青年使用的分析学基础》.艾格纳丝的课本表现出莱布尼茨及其跟随者的影响而不是牛顿的.因此它用的是微分和无穷小的语言而不是流数.(非常有趣地,英文译者把所有的 dx 都替换成 $\dot{x}$,尽管他经常保留"微分"一词而不是替换为"流数".)课本解释概念清楚并提供众多的例子.因此,在她的关于极大值和极小值的一节中,艾格纳丝提出这样的问题,即在某点分割某线使得一段的长度与另一段的平方的乘积最大,以及求过某矩形的一个顶点且交两对边的延长线的最短的线段.她甚至展示了如何求在由 $a(dy/y)=dx$ 所定义的对数曲线上的曲率最大的点.

对于艾格纳丝,如同对于约翰·伯努利一样,积分学是微分学的逆,亦即,从一个给定的微分的表达式确定使这表达式是微分的那个量的方法.记号 $\int y dx$ 意味着一个反导数.但是被积表达式的记号,那个 ydx 表示一个无穷小矩形的面积,实际上作为一种马后炮引导她去注意曲线下方的面积可以用这种同样的反过程计算.

艾格纳丝在处理对数(指数)曲线中是特别彻底的.她注意到积分的通常规则导致从 $dx=ay^{-1}dy$ 到 $x=ay^{-1+1}/(-1+1)$ 或 $ay^0/0$,这"什么也没告诉我们."于是她用另外的方式去处理这曲线.她首先表明随着横坐标算术地增加而纵坐标几何地增加的曲线具有微分方程 $dx=ay^{-1}dy$,此后这个式子可以利用适当的无穷级数进行计算.她也展示了如何求这曲线下方的在有限区间上和在从一个固定的横坐标 x 向左延伸的无穷区间上的面积.她计算这个"非正常积分"(今天被写为 $\int_{-\infty}^{x} e^{t/a}dt$)得到 ay,其中 y 是相应于 x 的纵坐标,并且展示了如何计算曲线绕 x-轴旋转生成的立体的体积.但是在18世纪上半叶的其他书籍中没有什么关于三角函数的内容.

如前面讨论的那另两本教科书的作者一样,艾格纳丝的名字奇怪地联系到在她的书中甚至并不源于她的一个小条目.作为在解析几何中的一个例子,她几何地描述了她所确定的方程为 $y=$

$\frac{a\sqrt{a-x}}{\sqrt{x}}$ 的一条曲线,一条早先已被命名为箕舌线(la versiera)的曲线,从拉丁意思"转弯"衍生而来.不幸地,versiera 一词也是意大利词汇 avversiera 的缩写,意思是"魔鬼之妻".因为英文译者把这个词翻译为"巫婆(witch)",这曲线自此也被叫作"艾格纳丝巫婆".

13.2.4 欧拉的《导论》

在此之前讨论的三种课本都是用本国语写的.可是,一系列拉丁文课本被证明对于未来是更加重要的.这些是欧拉的著作,有《Introductio in Analysin Infinitorum》(无穷分析导论,1748)两卷、《Institutiones Calculi Differentialis》(微分学原理,1755)和《Institutiones Calculi Integralis》(积分学原理,1768—1770).

《导论》,作为欧拉的"预微积分"课本,是发展"为微积分所绝对需要"的那些主题使得读者"几乎觉察不到地变得熟悉无穷的概念"[17] 的一个尝试.并且因为分析学与函数相联系,欧拉从它们的定义开始他的著作:"一个变量的**函数**是用任何方式由这个变量和数字或常量所构成的一个解析表达."[18] 关于欧拉的定义要注意的第一点是"函数"一词意思是一个"解析表达",亦即,一个公式.第二是他的对于如何构成这些公式的陈述,"用任何方式",只能通过考虑他的进一步的讨论来理解.对于欧拉有两个基本的函数类,代数的和超越的.前者是由变量和常数通过加、减、乘、除、乘幂、开根号,以及一个方程的解所构成的.后者是那些通过指数、对数,以及更一般地,通过积分而定义的.因为积分不能在预微积分著作中讨论,在《导论》中所讨论的超越函数限制在三角、指数和对数函数的特殊类型中.

欧拉讨论函数时的一个重要工具是幂级数.欧拉深信,或许除了在孤立点上,任何函数都可以用幂级数来表示,但是他没有给出证明.然而,他试图通过展示如何将任意的代数函数以及不同的超越函数展成幂级数来使读者相信这个事实.他的对于代数函数的方法不是新的,是对于牛顿法利用除法(在有理函数情形)和二项式定理(对于可用任何幂次的项来表示的函数)的一种组合.并且没有讨论收敛性.

可是,《导论》的具有最大影响的几节处理了指数、对数和三角函数,因为正是在这部分欧拉介绍了符号和概念,使得在早先课本中的这种函数的所有的讨论都作废.关于这些函数的所有的现代的处理,在某种意义上都是来自于欧拉.这样欧拉像幂函数一样来定义指数函数,在其中指数是变化的,然后,作为这样做的第一人,以此定义对数.就是说,若 $a^z=y$,欧拉定义 z 为以 a 为底的 y 的对数.对数函数的基本性质就从指数函数的性质得到了.

欧拉通过利用二项式定理对于任意底数的指数函数和对数函数发展了幂级数.他的技巧将"无穷小的"和"无穷大的"数都进行了重要的应用,在今天应用这些概念是令人皱眉的.然而,欧拉极少犯错误.例如,他注意到因 $a^0=1$,得到 $a^\omega=1+\psi$,这里 ω 和 ψ 都是无穷小的,因此 ψ 一定是某倍的 ω,依赖于 a,并且

$$a^\omega = 1 + k\omega \quad \text{或} \quad \omega = \log_a(1+k\omega).$$

欧拉接着注意到对于任意的 j,$a^{j\omega}=(1+k\omega)^j$,用二项式展开右边,有

$$a^{j\omega} = 1 + \frac{j}{1}k\omega + \frac{j(j-1)}{1\cdot 2}k^2\omega^2 + \frac{j(j-1)(j-2)}{1\cdot 2\cdot 3}k^3\omega^3 + \cdots.$$

若取 j 等于 z/ω,其中 z 是有限的,则 j 是无限大且 $\omega = z/j$.级数现在变成

$$a^z = 1 + \frac{1}{1}kz + \frac{1(j-1)}{1\cdot 2j}k^2z^2 + \frac{1(j-1)(j-2)}{1\cdot 2j\cdot 3j}k^3z^3 + \cdots.$$

因为 j 是无限大,$(j-n)/j=1$ 对任意正整数 n 成立. 展式于是化简成级数

$$a^z = 1 + \frac{kz}{1} + \frac{k^2z^2}{1\cdot 2} + \frac{k^3z^3}{1\cdot 2\cdot 3} + \cdots,$$

这里 k 依赖于底 a. 欧拉也注意到方程 $\omega = \log_a(1+k\omega)$ 蕴涵着,若 $(1+k\omega)^j = 1+x$,则 $\log_a(1+x) = j\omega$. 然后因为 $k\omega = (1+x)^{1/j}-1$,得到

$$\log_a(1+x) = \frac{j}{k}(1+x)^{1/j} - \frac{j}{k}.$$

二项式定理的另一灵巧的运用最终使他得到级数

$$\log_a(1+x) = \frac{1}{k}\left(\frac{x}{1} - \frac{x^2}{2} + \frac{x^3}{3} - \cdots\right).$$

取 $k=1$ 或者等价地取 $a=\mathrm{e}$,给出 e^z 和 $\ln z$ 的标准的幂级数.

欧拉的"从圆产生的超越量"的论述是第一本在讨论三角函数时对待这些量是作为函数具有数值而不是作为在固定半径的圆内的线段的教材. 实际上,欧拉并没有给出正弦和余弦的新的定义,他只是注意到他将总是考虑以半径为 1 的圆的方式定义的一段圆弧 z 的正弦和余弦. 正弦和余弦的包括加法和周期性质在内的所有基本性质被假定是已知的,尽管欧拉确实推出了一些相对复杂的恒等式. 更重要的是欧拉通过运用二项式定理和复数推出了正弦和余弦的幂级数. 从容易推出的恒等式,$(\cos z \pm \mathrm{i}\sin z)^n = \cos nz \pm \mathrm{i}\sin nz$,欧拉得出结论

$$\cos nz = \frac{(\cos z + \mathrm{i}\sin z)^n + (\cos z - \mathrm{i}\sin z)^n}{2},$$

以及,通过展开右端,有

$$\cos nz = (\cos z)^n - \frac{n(n-1)}{1\cdot 2}(\cos z)^{n-2}(\sin z)^2 + \frac{n(n-1)(n-2)(n-3)}{1\cdot 2\cdot 3\cdot 4}(\cos z)^{n-4}(\sin z)^4 + \cdots.$$

再令 z 为无限小,n 无限大,并且 $nz=v$ 有限,可以从 $\sin z = z$ 和 $\cos z = 1$ 得到

$$\cos v = 1 - \frac{v^2}{1\cdot 2} + \frac{v^4}{1\cdot 2\cdot 3\cdot 4} - \cdots.$$

《导论》的第一卷的其余部分包括了很多的关于无限过程的其他内容,包括无穷级数和无穷乘积. 例如. 欧拉展示如何通过正弦函数的性质确定无穷和 $\sum_{n=1}^{\infty}\frac{1}{n^{2k}}$,他还考虑等式

$$\prod_p \frac{1}{1-\frac{1}{p^n}} = \sum_m \frac{1}{m^n},$$

这里乘积取在所有的素数上而和取在所有的正整数上. 这个乘积及这个和,都被推广到 n 是任意复数 s 的情形,今天被称为变量 s 的黎曼 Zeta 函数,它的研究已产生很多新数学. 双曲函数也是以乘积和因子的形式出现,尽管它们没有被命名. 因此欧拉表示

$$\frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} = \frac{x}{2} + \frac{x^3}{1\cdot 2\cdot 3} + \frac{x^5}{1\cdot 2\cdot 3\cdot 4\cdot 5} + \cdots,$$

可以被因式分解成

$$\frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} = x\left(1+\frac{x^2}{\pi^2}\right)\left(1+\frac{x^2}{4\pi^2}\right)\left(1+\frac{x^2}{9\pi^2}\right)\cdots.$$

类似地

$$\frac{e^x + e^{-x}}{2} = 1 + \frac{x^2}{1 \cdot 2} + \frac{x^4}{1 \cdot 2 \cdot 3 \cdot 4} + \cdots = \left(\frac{1 + 4x^2}{\pi^2}\right)\left(\frac{1 + 4x^2}{9\pi^2}\right)\left(\frac{1 + 4x^2}{25\pi^2}\right)\cdots.$$

约翰·海因里奇·朗贝特(1728—1777)在 1768 年给这两个函数引进了名字双曲正弦($\sinh x$)和双曲余弦($\cosh x$),并给出了这两个函数与普通的正弦和余弦函数的类比.

13.2.5　欧拉的微分学

尽管《导论》的第一卷的大部分是关于级数的,(第二卷,关于解析几何将在第十四章里论述.)欧拉却是为了微积分学的代数需要而考虑这一内容的.他在 1755 年的他自己的《微分学原理》中讨论微积分学本身.那部著作从他的微分学的定义开始:"(它)是确定函数得到的无穷小增量与变量的那些无穷小增量之比的方法,它们是变量的函数."[19] 欧拉已经在《导论》中给出"函数"的一个定义,但是他在这里对它有少许一般化:"然而,当量以这样的方式依赖于其他的量,即(前者)随着(后者)改变而改变自身,则(前者)叫做(后者的)函数;这是一个很综合性的概念,在它自身当中包括了所有的一个量可能被其他的量所确定的方式."[20] 因此欧拉不再要求一个函数是一个"解析表达式".这个变化的理由或许和在 13.4 节中讨论的弦振动问题的争议有关.虽然欧拉熟知微分学在几何上的许多运用,但是他仍然坚持他的这本著作是一本纯粹的分析学的著作.如此,就没有任何的图示.

因为微分学必然与"无穷小增量"之比有关,欧拉从增量的一个一般讨论开始,也就是说,从有限差分开始.给定变量值的一个序列,即 $x, x + \omega, x + 2\omega, \cdots$,以及相应的函数值的一个序列 $y, y', y'', \cdots$,欧拉考虑各种不同的有限差分的序列.一阶差分是 $\Delta y = y' - y, \Delta y' = y'' - y', \Delta y'' = y''' - y'', \cdots$;二阶差分为 $\Delta\Delta y = \Delta y' - \Delta y, \Delta\Delta y' = \Delta y'' - \Delta y', \cdots$;三阶和更高阶差分均可类似定义.例如,如果 $y = x^2$,那么 $y' = (x + w)^2$ 且 $\Delta y = 2\omega x + \omega^2, \Delta\Delta y = 2\omega^2$,而三阶和更高阶差分均为零.利用各种技巧包括使用级数展开,欧拉可以计算所有的标准的初等函数的差分.进一步地,利用和 $\sum$ 来表示 Δ 运算的逆,他还给此运算推出各种不同的公式.这样,因为 $\Delta x = \omega$,就得到 $\sum \omega = x$ 和 $\sum 1 = x/\omega$.类似地,因为 $\Delta x^2 = 2\omega x + \omega^2$,推出 $\sum(2\omega x + \omega^2) = x^2$ 以及

$$\sum x = \frac{x^2}{2\omega} - \sum \frac{\omega}{2} = \frac{x^2}{2\omega} - \frac{x}{2}.$$

欧拉然后容易从 Δ 的相应的规则发展 $\sum$ 的规则.并不是讨论有限差分的规则,而却是讨论欧拉的微分的规则将更为有用.

"无穷小的分析学……不是别的而是差分法的一个特殊情形……当起先被假设为有限的差分取作无穷小时产生."[21] 从欧拉计算这些无穷小量即微分的规则产生了微分学的标准的公式.例如,若 $y = x^n$,则微分式

$$y' = (x + dx)^n = x^n + nx^{n-1}dx + \frac{n(n-1)}{1 \cdot 2}x^{n-2}dx^2 + \cdots,$$

因此

$$dy = y' - y = nx^{n-1}dx + \frac{n(n-1)}{1 \cdot 2}n^{n-2}dx^2 + \cdots.$$

"但是在表达式中第二项以及顺次的其他各项和第一项相比较将等于零."[22] 因而 $d(x^n) = mx^{n-1}dx$.这里应该注意的是欧拉要使他的论证不仅应用于 x 的正整数幂函数,而且应用于任意幂函数.最终二项式定理应用于所有幂函数.因此的展开不是必须表示一个有限和;它也可以表示一个无穷级数.欧拉因此立即注意到 $d\left(\frac{1}{x^m}\right) = -\frac{m dx}{x^{m+1}}$,以及更一般地,

$$d(x^{\mu/\nu}) = (\mu/\nu)x^{(\mu-\nu)/\nu}dx.$$

欧拉没有给出现代的链式法则的明确表述,但是当出现需要时他就处理特殊情形.因此,如果 p 是 x 的函数且 dp 是它的微分,则 $d(p^n) = np^{n-1}dp$.欧拉的乘积法则的推导和莱布尼茨的本质上相同,但是他的商的法则的推导是更加本原的.他把 $1/(q + dq)$ 展开成幂级数:

$$\frac{1}{q + dq} = \frac{1}{q}\left(1 - \frac{dq}{q} + \frac{dq^2}{q^2} - \cdots\right),$$

忽略高阶项,然后写作

$$\frac{p + dp}{q + dq} = (p + dp)\left(\frac{1}{q} - \frac{dq}{q^2}\right) = \frac{p}{q} - \frac{pdq}{q^2} + \frac{dp}{q} - \frac{dpdq}{q^2}.$$

由于相应于一阶微分的二阶微分 $dpdq$ 等于零,得到

$$d\left(\frac{p}{q}\right) = \frac{p + dp}{q + dq} - \frac{p}{q} = \frac{dp}{q} - \frac{pdq}{q^2} = \frac{qdp - pdq}{q^2}.$$

对数的微分需要用到在《导论》中得到的幂级数.若 $y = \ln x$,则

$$dy = \ln(x + dx) - \ln(x) = \ln\left(1 + \frac{dx}{x}\right) = \frac{dx}{x} - \frac{dx^2}{2x^2} + \frac{dx^3}{3x^3} - \cdots.$$

舍去高阶微分后欧拉立即得到公式.对反正弦函数的方法是通过复数.将 $y = \arcsin x$ 代入公式 $e^{iy} = \cos y + i\sin y$ 得到 $e^{iy} = \sqrt{1 - x^2} + ix$.推出 $y = \frac{1}{i}\ln(\sqrt{1 - x^2} + ix)$,以及因此有

$$dy = d(\arcsin x) = \frac{1}{i}\frac{1}{\sqrt{1 - x^2} + ix}\left(\frac{-x}{\sqrt{1 - x^2}} + i\right)dx = \frac{dx}{\sqrt{1 - x^2}}.$$

最后,欧拉对正弦函数的微分的推导从计算 $d(\sin x) = \sin(x + dx) - \sin x = \sin x\cos dx + \cos x \sin dx - \sin x$ 开始.欧拉然后回顾正弦和余弦的级数展开,再舍去高阶项,注意 $\cos dx = 1$ 和 $\sin dx = dx$.得到所要求的 $d(\sin x) = \cos x\, dx$.

欧拉关于二元或多元函数的微分的一章没有记载在这个概念的早期发展中的所作出的任何努力.他仅仅提到,在处理这样的函数时,变量可以独立地变化.这章中的中心思想,如在一元函数的情形,是关于微分和微分系数(我们的导数或偏导数)的.欧拉主要通过利用例子表明,若 V 是 x 和 y 两个变量的函数,则 dV,由 x 至 $x + dx$ 和 y 至 $y + dy$ 的变化而产生的 V 的变化,可以表达成 $dV = pdx + qdy$,这里 p、q 是将 y、x 分别作为常数产生的微分系数.计算 p 或 q 自然没有困难,因为用现代的形式,$p = \partial V/\partial x$ 且 $q = \partial V/\partial y$.只是运用已经得到的法则,将一个或另一个变量看作常数.欧拉进一步用含有微分的代数论证表明"混合偏导数"是相等的.

欧拉的课本有很多其他特点,包括对微分方程的一个介绍,在其中他展示了如何由一个给定的二元方程生成这样的方程,对泰勒级数的一个讨论,关于将函数转化成幂级数的各种不同的方法的一章,求各种不同的级数的和的一个广泛的讨论,包括整数的各种不同的幂的和,以及数值求解方程的根的种种方法.但是,这里讨论的其余部分将集中于欧拉的关于求极大值和极小值的两章.回顾到课本中没有任何图示并且因此没有具有极大值或极小值的曲线.每件事都分析地完成.但是正是通过区分一个绝对极大值,即一个比函数的任何其他值都大的值;和一个局部极大值,即一个 y 在 $x = f$ 处的取值,它比在两侧"靠近" f 的 x 的任何其他的 y 值都大,欧拉开始了讨论.但是他没有讨论"靠近"是什么确切的含义.

欧拉以一阶和二阶导数的方式,通过利用麦克劳林级数,用一种本质上与麦克劳林的相同的途径,得到一个函数在 $x = \alpha$ 处具有极大值或极小值的基本准则.但是欧拉比他的苏格兰前辈给出了更多的例子,并且经常寻求推广.因此,在考察了几个特殊的多项式的最大值和最小值之后,他较详

细地考察了一个任意的多项式 $y = x^n + Ax^{n-1} + Bx^{n-2} + \cdots + D$.在处理了几个有理函数的情形之后,他考虑了更一般的有理函数$\frac{(\alpha+\beta x)^m}{(\gamma+\delta x)^n}$.在讨论了缺少在0附近的关于 $x^{2/3}$ 的一个幂级数,以及因此需要将极大值或极小值的某些不同的准则公式化之后,他研究更一般的情形 $x^{2pz/(2q-1)}$.欧拉的大多数例子是代数方程的,但是他以利用超越函数的几个例子作为结束,包括函数和,这两者都需要做仔细的数值工作去得到关于极值的准确的解.

对二元函数 V,欧拉开始于考虑形如 $X + Y$ 的特殊情形的函数,其中 X 仅是 x 的函数、Y 仅是 y 的.在这个情形中,一对值(x_0, y_0)使得 x_0 是 X 的极大值且 y_0 是 Y 的极大值明显地给出 $X + Y$ 的极大值.对于更一般的情形,欧拉认识到,通过交替保持每个变量为常量,V 能出现极值仅当微分 $P\mathrm{d}x + Q\mathrm{d}y = 0$ 时,因此仅当 $P = \partial V/\partial x = 0$ 和 $Q = \partial V/\partial y = 0$ 都成立时.确定是否一点(x_0, y_0)当在此处两个一阶偏导数都为零时产生极大值、极小值或者两者都不是的问题是更加困难的,并且事实上,欧拉没能给出完整的结果.他断言到,若$\partial^2 V/\partial x^2$ 和$\partial^2 V/\partial y^2$ 在点(x_0, y_0)均为正,则函数 V 在该点有极小值;若它们均为负,则有极大值.欧拉举出几个例子来说明这个方法,包括 $V = x^3 + y^2 - 3xy + (3/2)x$.他的准则蕴涵当 $x = 1, y = 3/2$ 和 $x = 1/2, y = 3/4$ 时有极小值.不幸的是,后者不是极小值点,却是个鞍点.

13.2.6　欧拉的积分学

欧拉在分析学中的三部曲的最后部分,《积分学原理》,以积分学的一个定义开始.它是从某量的给定的微分关系中求出量本身的方法.也就是说,对于欧拉,如同对于艾格纳丝和约翰·伯努利一样,积分法是微分法的逆,而不是面积的确定法.因此著作的第一部分处理对各种类型的函数的积分(求反导数)的技巧,而课本的其余部分处理微分方程的解法.欧拉以这样的结果开始

$$\int ax^n \mathrm{d}x = \frac{a}{n+1}x^{n+1} + C,$$

对于 $n \neq -1$,以及

$$\int \frac{a\mathrm{d}x}{x} = a\ln x + C = \ln cx^a,$$

然后对部分分式技巧进行了详细的讨论使得可以对有理函数积分.他讨论了对于含有平方根的函数的积分的各种代换方法,虽然没有现代的三角代换法.有一章讨论通过使用无穷级数的积分法,它是牛顿喜欢的技巧,而另一章讨论分部积分法,尤其是在涉及对数和指数函数的情形,第三个则考虑三角函数的幂的积分的简约公式.欧拉甚至使用代换

$$\cos\phi = \frac{1-x^2}{1+x^2}, \quad \sin\phi = \frac{2x}{1+x^2}$$

将含有正弦和余弦函数的有理函数转化为普通的有理函数.

课本的大部分讨论解微分方程的方法.欧拉通过分离变量法求解一般的一阶线性方程 $\mathrm{d}y + Py\mathrm{d}x = Q\mathrm{d}x$(或用现代方式,$y' + Py = Q$)得到通解

$$y = \mathrm{e}^{-\int p\mathrm{d}x}\int \mathrm{e}^{\int p\mathrm{d}x}Q\mathrm{d}x,$$

这个结果他在1734年已证明.利用克莱罗的1739年的思想,他展示了在$\partial P/\partial y = \partial Q/\partial x$的"恰当"情形下如何对 $P\mathrm{d}x + Q\mathrm{d}y$ 积分.他讨论了在 $P\mathrm{d}x + Q\mathrm{d}y$ 不是恰当的情形下如何找积分因子,这是个他和克莱罗也早已发现的方法.他讨论了二阶和更高阶微分方程的各种情形,包括常系数线性的情形.最后,欧拉以偏微分方程的讨论结束此书,一个最重要的例子将在13.4中讨论.但是,即使考

虑微分方程的原来的动机大部分来自于物理的问题,但欧拉在课本中从未提到它们.

像《微分学》和《导论》一样,《积分学》是纯分析学的课本,以至于欧拉甚至没有讨论几何应用.这个事实或许可以解释在欧拉的著作与现代微积分学课本之间的重要差别.因此在《微分学》中,没有切线和法线、没有切平面,没有曲率的研究——所有这些主题欧拉在1740年都完全熟谙,但是它们仅仅出现于他的几何学著作中.甚至更令人惊奇地,在《积分学》中没有面积的计算,也没有关于曲线长度、或立体的体积、或表面积的任何内容.随后并没有出现现代工作的核心——微积分基本定理.甚至没有定积分的计算.欧拉当然非常擅长用反导数求面积,实际上他在各种论文中用这一思想.另一方面,由于在他的著作中没有出现将曲线下的面积作为函数的清晰的观念,故他没有考虑这种面积函数的导数.

不管现代读者在欧拉的微积分课本中可能发现的差距,它们到18世纪末最终被证明在对欧拉及其前辈所发展的内容提出一个结构和一个清楚的解释中是有影响力的.18世纪下半叶的所有数学家不断使用欧拉的著作.可是,到下个世纪的早期,学生的需求开始改变.正是在法国大革命的变革之后新学生进入科学领域,刺激撰写出版了许多新教材,这些教材取代了欧拉的并成为今天的课本的直接祖先.

13.3 重积分

虽然欧拉在他的《积分论》中没有讨论重积分,但是他参与了这个课题的重要概念的发展,这个课题始于莱布尼茨对文森佐·维维安尼(Vincenzo Viviani)(1622—1703)在1692年4月4日的挑战问题的解答.维维安尼,隐藏在变位字D.Pio Lisci Pusillo Geometra(意为Postremo Galileo Discipulo(即伽利略的关门弟子))之后,提出在半球面上确定四个相等的"窗口"使得表面的剩余部分在面积上等于用直尺和圆规可画出的区域的面积的问题.莱布尼茨在接到问题的当天1692年5月27日解决了这个问题.在解决它时,他必须计算在半球面上的不同区域的面积,为此他将含有两个微分的乘积的表达式积分,通过对第一个变量先积分,并将第二个看作常数,然后对第二个积分.

这个问题和类似的问题稍后被伯努利和洛必达在其它问题中解决了,但是直到1731年计算某区域的体积及其边界曲面的面积的系统的努力才由克莱罗在他的《关于二重曲率的曲线的研究》中发表.尽管克莱罗演示了曲面通常可用一个三元方程表示,他最常考虑的却是由一个坐标平面内的一条曲线生成的圆柱面.于是为了计算夹在由 $y = f(x)$ 和 $z = g(x)$ 给出的两个圆柱之间的区域的体积,他表明体积元素为 $\mathrm{d}x\int z\mathrm{d}y$,然后用方程将 z 和 $\mathrm{d}y$ 改写成 x 的形式使得可以对 $z\mathrm{d}y$ 积分.鉴于体积元素现在完全以 x 的形式给出,他可以再积分计算要求的体积.类似地,他将表面积元素表示成 $\mathrm{d}x\int\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$ 并进行类似的计算.

1760年在关于变分的微积分学的工作中,约瑟夫·路易斯·拉格朗日(1736—1813)也不得不讨论体积和表面积.拉格朗日仅对体积写出 $\iint z\mathrm{d}x\mathrm{d}y$ 并对表面积写出 $\iint \mathrm{d}x\mathrm{d}y\sqrt{1 + P^2 + Q^2}$,其中曲面的方程是 $z = f(x,y)$ 并且这里 $\mathrm{d}z = P\mathrm{d}x + Q\mathrm{d}y$.这些记号不带有更多的讨论出现在他于1750年给欧拉的信中和他早期的论文中,虽然他没有注意到二重积分符号表示着两个积分必须依次进行.

人物小传	阿列克塞 · 克劳德 · 克莱罗(1713—1765)(Alexis Claude Clairaut)
	克莱罗出生于巴黎,是一名天才儿童.他 10 岁掌握洛必达的《无穷小分析》,并且两年后在巴黎科学院宣读论文.从 13 岁开始的关于曲线的研究使他 1731 年出书并且使他最终在 18 岁被选入科学院.克莱罗不久转到天体力学并且后来转到教育学.他在以前的领域中的五部著作被证明是很有影响力的,并且他的关于几何学和代数学的两本课本,是介绍教授这些科目的一种历史的或"自然的"方法的尝试,第一本将在第 14 章讨论.

13.3.1 二重积分的概念

只是在 1769 年的一篇论文中欧拉才给出了二重积分概念的第一个详细的解释.欧拉从推广他的积分作为反导数的概念开始.因此$\iint Z\mathrm{d}x\mathrm{d}y$ 就意味着一个两个变量的函数,其中两个变量当第一次对 x 第二次对 y 被两次微分时作为微分给出 $Z\mathrm{d}x\mathrm{d}y$.例如,欧拉注意到$\iint a\mathrm{d}x\mathrm{d}y = axy + X + Y$,其中 X 是 x 的函数 Y 且是 y 的函数.一个更复杂的例子为

$$\iint \frac{\mathrm{d}x\mathrm{d}y}{x^2 + y^2}.$$

对两个变量之一进行第一次积分均可.按每一方法积分,欧拉得到值

$$\int \frac{\mathrm{d}x}{x}\arctan\frac{y}{x} + X \quad \text{和} \quad \int \frac{\mathrm{d}y}{y}\arctan\frac{x}{y} + Y.$$

因为对此两个积分中每个进行第二次积分的惟一方法是通过将被积函数写成幂级数,欧拉这样作了并且表明了两种积分法都导出相同的最后结果

$$\iint \frac{\mathrm{d}x\mathrm{d}y}{x^2 + y^2} = X + Y - \frac{y}{x} + \frac{y^3}{9x^3} - \frac{y^5}{25x^5} + \cdots.$$

然后给予二重积分作为二重反导数的思想,欧拉把通过一重积分求面积的观念推广到利用二重积分求体积.他的基本思想,像莱布尼茨一样,是对一个变量先积分,保持另一个为常数,然后对付第二个.他的第一个例子是求半径为 a、其方程为 $z = \sqrt{a^2 - x^2 - y^2}$ 的八分之一球面下方的体积.在 x-y 平面的圆的第一象限中取面积元素 $\mathrm{d}x\mathrm{d}y$,欧拉注意到在这无穷小矩形上方的柱体体积为 $\mathrm{d}x\mathrm{d}y\sqrt{a^2 - x^2 - y^2}$(图 13.11).为确定这个体积,欧拉先对 y 积分,保持 x 为常量,得到

$$\left[\frac{1}{2}y\sqrt{a^2 - x^2 - y^2} + \frac{1}{2}(a^2 - x^2)\arcsin\frac{y}{\sqrt{a^2 - x^2}}\right]\mathrm{d}x,$$

作为在宽为 $\mathrm{d}x$ 且长为 $\mathrm{d}y$ 的矩形之上在球面片之下的体积.用 $\sqrt{a^2 - x^2}$ 代替 y,欧拉计算出对于那个 y 值的相同的片的体积是$\frac{\pi}{4}(a^2 - x^2)\mathrm{d}x$.对 x 积分后得到$\frac{\pi}{4}\left(a^2x - \frac{1}{3}x^3\right)$为从 y 轴到 x 的体积,并且用 a 代替 x 后得出八分之一球的总体积为$\frac{\pi}{6}a^3$,以及整个球的体积为$\frac{4\pi}{3}a^3$.

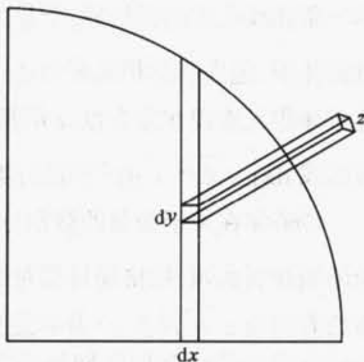


图 13.11 球的体积.

在进一步表明如何计算由在上面用球面和在下面用各种

平面区域所围成的立体区域的体积之后，欧拉注意到二重积分也可以被用来计算表面积，大概已知拉格朗日早先给出的一般公式，他没有过多讨论便给出了球面的表面积元素为

$$\frac{a\mathrm{d}x\mathrm{d}y}{\sqrt{a^2-x^2-y^2}}.$$

进一步地，他还注意到在区域 A 上的 $\iint \mathrm{d}x\mathrm{d}y$ 准确地等于 A 的面积.

13.3.2 重积分的变量替换

欧拉关于二重积分的论文的最有趣的部分是他对于当变量被替换时二重积分发生什么的讨论. 换句话说，若 x 和 y 被给定为两个新变量 t 和 v 的函数，欧拉想要确定如何将 $Z\mathrm{d}x\mathrm{d}y$ 的积分变换成为具有面积元素 $\mathrm{d}t\mathrm{d}v$ 的新积分.

欧拉认识到，若给定的变量替换是平移接着旋转，即若

$$x = a + mt + v\sqrt{1-m^2},$$

$$y = b + t\sqrt{1-m^2} - mv,$$

这里 m 是旋转角 θ 的余弦，则面积元素 $\mathrm{d}x\mathrm{d}y$ 和 $\mathrm{d}t\mathrm{d}v$ 应该是相等的. 但是当他进行显然正常的计算

$$\mathrm{d}x = m\mathrm{d}t + \mathrm{d}v\sqrt{1-m^2},$$

$$\mathrm{d}y = \mathrm{d}t\sqrt{1-m^2} - m\mathrm{d}v.$$

并将两个方程相乘时，他得到

$$\mathrm{d}x\mathrm{d}y = m\sqrt{1-m^2}\mathrm{d}t^2 + (1-2m^2)\mathrm{d}t\mathrm{d}v - m\sqrt{1-m^2}\mathrm{d}v^2,$$

是一个他认为显然是错误的结果. 明显地，若 t、v 通过更复杂的变换与 x、y 相联系，类似的计算还会是错误的. 之后，欧拉所要作的是发展在上述情况下给出 $\mathrm{d}x\mathrm{d}y = \mathrm{d}t\mathrm{d}v$ 和在更一般的情形下给出 $\mathrm{d}x\mathrm{d}y = W\mathrm{d}t\mathrm{d}v$ 的一种方法，其中 W 是 t 和 v 的某个函数.

欧拉的思想是一次处理一个变量，正如在二重积分中所作的. 因此，他先引进新变量 v 使得 y 是 x 和 v 的函数. 然后 $\mathrm{d}y = P\mathrm{d}x + Q\mathrm{d}v$. 假设 x 固定，欧拉得到

$$\mathrm{d}y = Q\mathrm{d}v \quad 和 \quad \iint \mathrm{d}x\mathrm{d}y = \iint Q\mathrm{d}x\mathrm{d}v = \int \mathrm{d}v\int Q\mathrm{d}x.$$

类似地，现在令 x 为 t 和 v 的函数使得 $\mathrm{d}x = R\mathrm{d}t + S\mathrm{d}v$ 且保持 v 为常数，他算得

$$\int \mathrm{d}v\int Q\mathrm{d}x = \int \mathrm{d}v\int QR\mathrm{d}t = \iint QR\mathrm{d}t\mathrm{d}v.$$

这样他对这个问题的初始的解答是 $\mathrm{d}x\mathrm{d}y = QR\mathrm{d}t\mathrm{d}v$，但是这个回答不能完全令人满意，因为 Q 可以依赖于 x，并且因为这个方法不对称. 欧拉因此继续考虑 y 为 t 和 v 的函数，并且重新计算 $\mathrm{d}y$:

$$\mathrm{d}y = P\mathrm{d}x + Q\mathrm{d}v = P(R\mathrm{d}t + S\mathrm{d}v) + Q\mathrm{d}v = PR\mathrm{d}t + (PS + Q)\mathrm{d}v.$$

因为 $\mathrm{d}y = T\mathrm{d}t + V\mathrm{d}v$ 也是成立的，得到 $PR = T$ 和 $PS + Q = V$ 或者 $QR = VR - ST$. 欧拉的最后解答是 $\mathrm{d}x\mathrm{d}y = (VR - ST)\mathrm{d}t\mathrm{d}v$. 他进一步注意到，事实上必须使用 $VR - ST$ 的绝对值，因为面积是正值. 欧拉的结论用现代符号表示为函数的二重积分就是

$$\iint f(x,y)\mathrm{d}x\mathrm{d}y = \iint f(x(t,v),y(t,v))\left|\frac{\partial x}{\partial t}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial t}\right|\mathrm{d}t\mathrm{d}v,$$

这里积分所取的区域是被在 (x,y) 和 (t,v) 之间的给定的函数关系所联系的.

正如 18 世纪很典型的数学证明，欧拉的论证是形式的. 拉格朗日使用类似的形式论证得到三维积分的变元公式. 两人都没有用到极限的概念或无穷小量近似或甚至担心是否在某点相关的导

数可能不存在.虽然尤其是欧拉通过运用这样的论证在发展新数学中取得了巨大的成功,但是缺乏在希腊几何学的模型上的公理基础困扰着他同时代的某些人.因此发展出一场关于微积分学的主要概念的恰当基础的争论,这场争论将在 13.5 节中介绍.

13.4　偏微分方程:波动方程

偏微分方程理论的开端发生在 18 世纪中叶,不仅有欧拉的最终记录在他的《积分论》中的工作,而且也有让·勒·隆德·达朗贝尔(1717—1783)(Jean Le Rond d'ALembert)和丹尼尔·伯努利的工作.我们在这里将仅讨论一种特殊类型的偏微分方程,波动方程,因为在引出方程的弦振动的课题上的争论不仅导致某些特殊的解法,而且导致函数概念的新理解.

人物小传	让·勒·隆德·达朗贝尔(1717—1783)(Jean Le Rond d'ALembert)
	达朗贝尔(图 13.12) 还是婴儿时就被他的刚放弃修女誓言并害怕报复的母亲,抛弃在巴黎的一座教堂的台阶上.他不久被一个贫穷的家庭收养,但是他富有的父亲给他提供丰厚的年金并且帮助他获得马萨林学院的录取,在这里他接受了古典文学艺术的教育. 虽然他 1737 年成为一名律师,但是他的真正兴趣在于他自学的一门学科 —— 数学.在发表了数篇论文特别是微分方程领域的论文之后,他 1741 年被巴黎科学院接受,并且很快成为欧洲最重要的数学家之一.他的著作不仅包括关于动力学和流体力学的主要论著,而且包括在 1750 年之后的目的是为了促进所有艺术和科学的基本原理的 28 卷本法国大百科全书的许多节.

13.4.1　达朗贝尔的工作



图 13.12　法国邮票上的达朗贝尔.

关于弦振动的课题的讨论开始于 1747 年达朗贝尔写的一篇论文,其中他提出对于处在振动中的绷紧的弦的形状的问题的一个解答.因为在弦上的点的位置随着横坐标和时间两者而变化,这个形状是由二元函数 $y = y(t, x)$ 确定的.达朗贝尔认为弦由无数个无穷小的质量构成并且然后使用牛顿定律得到一个对 y 的偏微分方程,现在称为**波动方程**,用现代的记号表示为

$$\frac{\partial^2 y}{\partial t^2} = c^2 \frac{\partial^2 y}{\partial x^2}.$$

达朗贝尔然后对 $c^2 = 1$ 的特殊情形求出方程的形如 $y = \Psi(t + x) + \Gamma(t - x)$ 的解,这里 Ψ 和 Γ 是任意的二次可微函数.如达朗贝尔指出"这个方程包含着无穷多条曲线."[23] 这个详尽的陈述导致了许多争论.

达朗贝尔自己首先讨论了当 $t = 0$ 时对每一 x, $y = 0$ 的情形,亦即,当 $t = 0$ 时弦在平衡位置.他进而要求对所有 t 在 $x = 0$ 和 $x = l$ 时 $y = 0$,即弦在区间$[0, l]$ 的两端被固定.第一个要求表明 $\Psi(x) + \Gamma(-x) = 0$,而第二个要求给出结果 $\Psi(t) + \Gamma(t) = 0$ 和 $\Psi(t + l) + \Gamma(t - l) = 0$. 因而得到 $\Gamma(t - x) = -\Psi(t - x)$ 或 $y = \Psi(t + x) - \Psi(t - x)$;得到 $\Psi(-x) = -\Gamma(x) = \Psi(x)$ 或 Ψ 是一个偶函数;并得到 $\Psi(t + l) = \Psi(t - l)$ 或 Ψ 是以 $2l$ 为周期的.进一步地,由于在$t = 0$ 点的初速度由 $\partial y/\partial t$ 给出,即 $v = \Psi'(x) - \Psi'(-x)$, 并且因为偶函数的导数为奇函数,达朗贝尔得出结论,"初速度的表达式 …….一定使得当化为级数时它只含有 x 的奇数次幂.否则 …… 问题

就会是不可能的."[24] 在不久后的一篇论文中,达朗贝尔把这个解法推广到弦的初始位置为 $y(0,x) = f(x)$ 和初速度为 $v(0,x) = g(x)$ 的情形.在这种情形,他得到的结果是,只有当 $f(x)$ 和 $g(x)$ 是周期为 $2l$ 的奇函数时解答才是可能的,并且为了使人可以对此进行运算,每个函数给定为一个二次可微解析表达式.那时,对达朗贝尔而言,一个函数恰如欧拉在他的《导论》中定义的那样.因此虽然 $f(x)$ 和 $g(x)$ 只是在区间$[0,l]$上给出,但是确定它们的函数 $y = \Psi(u)$ 必须自己是被给成一个对所有 u 值定义的"方程".其他类型的函数不能作为一个物理问题的解答出现.正如达朗贝尔三年后在另一篇论文中所得出的结论,"在任何其他情况下,这个问题都不能解决,至少通过我的方法是如此的,并且我不能肯定是否它没有超出已知的分析学的能力范围…….在这种假定下,我们仅在弦振动的所有不同的形态都可以通过一个相同的方程包含的情形下求解这个问题.对我来说,似乎在任何其他情况下将不能给出 y 的一般形式."[25]

13.4.2 欧拉和连续函数

在达朗贝尔的第一篇论文的两年以后,欧拉对相同的问题发表了他自己的解答,虽然用有些不同的推导却得出与达朗贝尔的相同的正式结果.但是欧拉和达朗贝尔的不同点在于允许什么类型的初始位置函数 f 上.首先,他宣布可以是任意的定义在$[0,l]$上的曲线,甚至不是由解析表达式确定的曲线.它可以是用手画的一条曲线.因此函数没有必要在每点都是可微的.第二,只是在初始区间上的定义才是重要的.可以通过在$[-l,0]$上用 $f(-x) = -f(x)$ 简单地定义它使曲线成为奇的和周期的,然后通过利用 $f(x \pm 2l) = f(x)$ 将它延拓到整个实数轴.最终,与达朗贝尔相反,欧拉推理到,就物理状态而言,弦的初始形状可以是任意的.即使有函数不可微的孤立点,仍然可以认为这条曲线为微分方程的解,因为欧拉认为在孤立点的行为与在区间上的函数的一般行为无关.

欧拉已经在他的《导论》的第二卷中定义了连续的和不连续的曲线的概念:"一条**连续**的曲线就是一条使它的性质可以被的一个单一函数所表示的曲线,若该曲线具有这样一种性质使得对它的不同部分……需要不同的 x 的函数来表达,也就是说,在一部分……由 x 的一个函数定义之后,然后需要另一个函数来表示(另)一部分……,于是我们称这样的曲线为**不连续的**…….这是因为这样的一条曲线不能被一个恒定的规律表示,而是由几个连续的部分构成."[26] 这样,在这场争论中达朗贝尔坚持初始条件只能用"连续的"曲线,而欧拉回应到曲线的不同部分之间没有必要按这样的连续规律连接.最终,欧拉改变了他的"函数"的定义以反映他在这种情况下的新观点,并且稍后他给达朗贝尔写到,对于这样的函数的考察"给我们开辟了分析学的一个全新领域."[27] 然而,达朗贝尔继续认为这样的"不连续的"曲线处在分析学的范围之外.

13.4.3 丹尼尔·伯努利和物理弦

欧拉和达朗贝尔,尽管他们以一般的"函数"的方式进行他们的分析,都在心中怀有正弦和余弦函数的例子.前者是奇的和周期的而后者是偶的和周期的.实际上,在1750年达朗贝尔通过分离变量法推出波动方程的解 $y = (A\cos Nt)(B\sin Nx)$,亦即,通过假设 $y = f(t)g(x)$ 然后微分.尽管如此,明确地提到正弦和余弦的结合试图使争论回到物理弦这一现实的,是争论的第三个参加者丹尼尔·伯努利,他在巴塞尔大学的职责包括医学、玄学和自然哲学,并且他的主要工作是在流体动力学和弹性力学中,在1753年的一篇论文中写到"达朗贝尔先生和欧拉先生的计算,当然囊括了分析法所可以具有的最深刻和最崇高的一切,……同时表明了不经过讨论问题任何综合检验地就被接受的一个抽象的分析法,容易令我们惊奇而不是使我们受到启发.就我来看似乎是,我们不经任

何计算而只须注意弦的简单振动的本性就可以预见这两位伟大的几何学家已经通过分析的头脑能进行的最棘手的和抽象的计算所发现的一切."[28]

伯努利对这个问题的更物理的解答是探索这样的观念,即弦振动潜在地表示无穷多音调,每个叠加在其他之上,每个单独用一条正弦曲线表示.尽管伯努利没有写出一般形式的结果,但是由此得出弦振动的位移可以表示为函数

$$y = \alpha \sin \frac{\pi x}{l} \cos \frac{\pi t}{l} + \beta \sin \frac{2\pi x}{l} \cos \frac{2\pi t}{l} + \gamma \sin \frac{3\pi x}{l} \cos \frac{3\pi t}{l} + \cdots$$

这里和是无穷的.欧拉和伯努利争论的初始位置函数可以用无穷和表示为

$$y(0, x) = \alpha \sin \frac{\pi x}{l} + \beta \sin \frac{2\pi x}{l} + \gamma \sin \frac{3\pi x}{l} + \cdots.$$

很有趣地,欧拉在 1750 年已经写出这些级数,可能仅想要以有限和来作为这个方程的可能解的一个例子.伯努利认为,用适当选取的常数 $\alpha, \beta, \gamma, \cdots$,后一个级数可以表示任意的初始位置函数 $f(x)$,但不能为他的观点的正确性给出数学的论证,后来他仅稍微写到,他的表示式提供无数个常数,它们可以用作调整曲线通过无数个特定点.他的观点受到欧拉的挑战,后者不仅不能看出任何确定这些系数的方法,而且认识到一个被表示成这样的三角级数的函数必定是周期的.不过通过这个论证,欧拉表现出自己羁绊于他原来的函数是什么的观点和他有助于帮助发展的更现代的观点之间.最终欧拉愿意允许定义在区间[0,1]上的任意曲线 $f(x)$ 被周期延拓到整个实数轴.但是这是一个可以被称作几何周期性的一个例子.它不考虑表示 f 的代数表达式.另一方面欧拉反对伯努利的理由是基于在整个实数轴上的三角函数本身的代数周期.欧拉只稍微察觉到现代的函数的定义域的概念,带着在定义域的不同部分函数可以表示为不同的表达式的一种伴随的可能性.

关于可接受为波动方程的解的函数类型的争论在接下来的几十年里由这三位数学家继续着,谁也不能被别人说服.虽然其他的数学家也加入到争论中,但直到 19 世纪早期才通过三角级数本性的全面分析获得解决.

13.5　微积分学的基础

18 世纪见证了微积分学技巧的广泛的发展.为想学的人们撰写了许多诠释微积分学的课本,出现了展示解决各种类型的物理问题的新的过程和方法的大量的论文.但是在某些人的心中有个关于这门学科的基础的不断的责问.大多数的数学家读过欧几里得的《原本》,认为它是应该如何做数学的典范.至此微积分学的核心过程仍然没有逻辑基础.一般地,专业人员自己并不很担心这一点.牛顿、莱布尼茨、欧拉和其他人对这个学科有着强烈的直觉,知道什么时候他们做什么是对的.即使用现代标准的眼光来看,这些伟大数学家也几乎没有做错过.尽管如此,他们自己作出的对他们的过程的基础的解释留下了某些想要的东西.



图 13.13　爱尔兰邮票上的伯克利.

13.5.1　乔治·伯克利的《分析学家》

对无穷小量和流数的最重要的批评是由爱尔兰哲学家毕晓普·乔治·伯克利(Bishop George Berkely)(1685—1753)(图 13.13)在 1734 年的题为《分析学家》(The Analyst)的小册子中作出的.这本著作是"给一位不信教

的数学家”的致词,一般推测是指天文学家艾德蒙·哈里(Edmond Halley),哈里资助了牛顿的《原理》的出版,并助之付印.伯克利推测地认为他是个不信教的人,因为他对一个共同的朋友劝说道基督教的教义是不可思议的.伯克利的《分析学家》的目的不是否定微积分学的用途或它的许多新的成果的有效性,而是表明数学家对他们所引起的过程没有有效的论证.

因此“流数的方法是通过用它来帮助现代数学家打开几何从而自然的秘密的普遍的钥匙”.可是,流数本身“被说成几乎是流动的量的增量(动差),它在最小的相等的时间极小差中产生;并且准确地说它是开始形成的最初比,或者它是渐进于零的增量的最后比 ….用动差我们不能理解有限的极小量 …[除了]仅仅有限量的开始形成的原理.”[29] 什么是“开始形成的原理”?伯克利指出即使“最微小的错误在数学中都不能被忽略”—— 来自牛顿自己的引语,实际求由开始形成的原理所确定的这些流数确实含有这类忽略.

伯努利通过分析 x^n 的流数的计算来说明他的观点:

> 在相同的时间里,x 通过流动变化为 $x+o$,x^n 变化为$(x+o)^n$,利用无穷级数的方法,
>
> $$x^n + nox^{n-1} + \frac{n^2-n}{2}o^2x^{n-2} + \cdots,$$
>
> 并且增量 o 和 $nox^{n-1} + \frac{n^2-n}{2}o^2x^{n-2} + \cdots$ 之比等于 1 和 $nx^{n-1}\frac{n^2-n}{2}ox^{n-2} + \cdots$ 之比.现在令增量等于零,最后的比值将等于 1 比 nx^{n-1}.但是似乎这个推理是不公正和不明确的.因为照它说的,令增量为零,亦即,令增量为无,或令没有增量,则前面假设增量是某种东西或有增量但被破坏了,然而至此那假设的结果,亦即,由此得到的表达式,却保留下来了.[30]

伯克利因此质问,如何能取一个非零的增量并用它进行计算,然后在最终却使它等于零.他进一步指出大陆的数学家做得不同.不是考虑“流动的量和流数,他们认为变有限量是通过持续地增加或减少无穷小量而增加或减少的”.这导致完全相同类型的问题.尤其是伯克利宣称他不能接受无穷小量.“但是要设想一部分这样的无穷小量仍会有比它更无穷小的量,而且通过无穷次地相乘的结果将永不会等于最微小的有限量,我猜想,这对任何人都是一个无限的困难.”[31] 因此二阶微分,类似地,流数的流数形成一个“模糊的奥秘.开始速度的开始速度,开始增加的开始增加,亦即,没有大小的东西,将它放在你所喜欢的场合,若我没有弄错,它的清楚的概念将是不可能被发现的.”[32] 由于哈里不能理解神学的争论,伯克利反击道“他能消化第二个或第三个流数,第二个或第三个差分,我认为不必对神的观点感到太拘谨.”[33]

13.5.2 麦克劳林对伯克利的应对

伯克利对微积分学基础的批评是有效的,关于一个值何时为零和何时不为零的问题甚至回溯到费马的工作,牛顿和莱布尼茨都不能较好地解决它.尽管如此,几位英国数学家出来投入到牛顿在伯克利攻击下的辩护中.最重要的回应是麦克劳林的《流数论》(Treatise of Fluxions).正如他在序言中写到,麦克劳林想要“通过最严格的形式的演绎从几个无懈可击的原理追随古人的方式推出[流数理论的]那些元素.”[34] 他提到他将不使用任何时间或空间的不可分的或无穷小部分来作为演绎的成分.“无穷小量的假设对几何这样一门学科是过于大胆的,”[35] 麦克劳林因此不得不认为有限长度和有限时间是他的基本元素,因为“没有比时间和空间的有限部分能被更清楚地接受的量了.”[36] 这些空间和时间决定了(平均)速度.但是因为流数理论所必需的基本概念是瞬时的速度,麦克劳林也尝试着给它下了一个定义:“在任何给定的时间间隔里一个变化的运动的速度不是由在给

定时间的那段间隔之后实际描述的空间来度量，而是由假如从那个间隔运动一致地继续就会描述的空间来度量."[37] 这个定义使人想起 14 世纪海泰斯布雷(Heytesbury) 给出的定义.可是,从现代的观点来看,麦克劳林给出了这样的定义,丢失了瞬时速度是随着时间区间趋近于零时的平均速度的极限的基本思想.无论怎样,给出变速度的定义之后,麦克劳林为了应用这个定义提出公理,然后着手用"古人的方式",每次利用二重归谬法,证明了大量的定理.

特别地,因为麦克劳林的目的之一是证明在牛顿论证中的无穷小量总可以被有限量代替,他演绎到即使微分三角形也可以被严格地推出:

命题 设 ET 为曲线 FE 在点 E 的切线,EI 与底边 AD 平行,设 IT 与纵轴 ED 平行,则底边、纵坐标和曲线的流数可以分别用线段 EI、IT 和 ET 度量(图 13.14).[38]

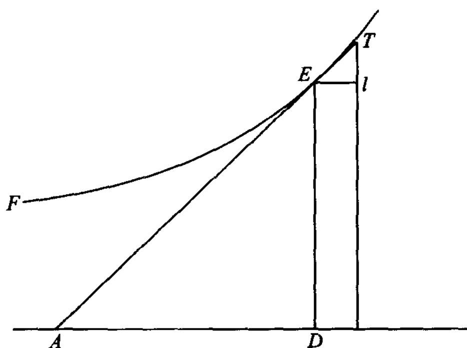


图 13.14 麦克劳林的微分三角形.

麦克劳林这样证明这个结果,首先指出若曲线是上凹的,则在底边的给定增量的时候纵坐标的增量比假如纵坐标的运动是一致时会生成的纵坐标的增量更大.为了证明后者的增加正比于纵坐标 DE 的流数,精确地等于 IT,他作出初始假设即它大于 IT 并且利用关于速度的他的公理证明一个矛盾的结果.一个类似的矛盾从相反的假设中得到.在曲线为下凹的情形整个证明则是重复的.从现代的观点来看,麦克劳林证明的问题在于他的切线的定义,这个概念被他同时期的人作为"自明地"普遍接受.可是,正符合于他对古人的方法的信念;麦克劳林提出的古代定义,即某曲线的一条切线是一条直线,它以这样的一种方式接触曲线,使得任何其他的直线都不能插在这条曲线和这条直线之间.

但是不管他的定义将瞬时速度作为基本概念并且不管他的切线的几何定义,麦克劳林熟知牛顿的"渐近于零的量的最终比"或"极限"的概念的应用.因而他记述了这个比,它是当两个增量减少直到为零时两个变量的有限瞬时增量彼此之间的各种比例的极限.他指出,为了发现这个极限必须先一般地确定增量的比,然后简化到最简单的形式使得一部分结果独立于增量本身的值.若假设增量"减少直到它们为零,"[39] 欲求的极限则欣然出现.例如,要求 x^2 的流数和 ax 的流数的比,麦克劳林计算了增量的比(当 x 增加到 $x+o$) 为 $2xo+o^2 : ao$ 或 $2x+o : a$."这个 $2x+o$ 与 a 的比当 o 减小时不断地减小,并且总是大于当 o 是任意的真实增量时 $2x$ 与 a 的比,但很明显它不断地接近于 $2x$ 与 a 的比作为它的极限."[40]

麦克劳林强烈地否定了伯克利的论点,即首先假设一个有限增量然后令这个增量为零的方法是矛盾的.实际上,他指出,这个方法允许确定当增量是有限值时增量的比以及确定比值怎样按增量变化.于是可以容易地确定当增量减小时比值趋近于什么极限.作为对伯克利的最后的反击,他甚至将切线定义为一个极限:"切线 …… 是这一条 …… 限制着所有通过接触点的割线的位置的直线,虽然严格地说它不是割线,(只是) 一个可以限制增量的变比值的比,虽然它不能被说成是任意

实增量的比."[41]

关于麦克劳林对微积分学的处理的问题,如同他许多同辈所说的那样,不是他不能严格地推导出这些规则,他是真正地按照"古人的方式"做它的.特别是他用了穷竭法和伴随它的归谬法论证.这样一种方法的运用给读者施加了很重的负担.例如,这本754页的著作的前590页没有任何流数符号.每一个新的思想通过非常冗长的几何推导得出.在18世纪很少有人愿意通读这些详尽的论证.麦克劳林自己认识到新微积分学的益处在于它能够以迅速而有效的形式解决旧的问题并且能够轻松地作出新的发现."但是古人的原理和严格的方法,它们至今维持着这个科学根据的完整性,如果到今天被抛弃的话,几何学家确定应该在哪里停止是困难的."[42] 尽管如此,虽然麦克劳林巨大的努力回答了伯克利的异议,18世纪的大多数的数学家对此并不欣赏.人们把自己看作是开创一片新天地而不是延续古人的方法.

13.5.3 欧拉和达朗贝尔

可是,即使在欧洲大陆这个方法的正当理由也是必需的.在欧拉的《微分学》中,他发展了这一思想,即包含在导数的计算中的比值实际上简单地是 $0:0$ 的形式.对于欧拉,无穷小量是实际上等于零的量,因为后者是惟一的比任意给定的量都小的量."从而没有像通常所认为的那么多的神秘隐藏在这个概念中."[43] 但是,虽然两个零以它们的差总为零的方式相等,欧拉坚持认为两个零的比值,依赖于正在变为零的量的起源,一定要在每个特定的情形下计算.作为例子,他指出 $0:0=2:1$ 是一个正确的命题,因为在等号每一端的第一个量是第二个量的两倍.实际上,那么 $0:0$ 完全可以等于任意的有限的比值.因此,"无穷小的计算 … 只不过是不同的无穷小量的几何比值的研究罢了."[44]

欧拉在这条主线上继续讨论,例如,进一步指出,"无穷小量同有限量相比较变为零,鉴于考虑有限量因而可以将其舍弃,"从而"无穷小的分析忽略数学的严格性的反对意见消失了 … 因为没有舍弃别的什么只是根本就没有什么."[45] 类似地,无穷小量 dx^2 相对于 dx 将变为零,因为 $(dx \pm dx^2)/dx = 1 \pm dx = 1$ 因而可以被忽略.

很有趣地,达朗贝尔在1754年写给《百科全书》的"微分学"一文中,合并了欧拉和麦克劳林的思想.他同意欧拉在看待 $0:0$ 上不存在错误,因为它实际上完全可以等于任意的量.但是微分学的中心思想是 dy/dx 当相关的量趋近于零时某个比值的极限."最精确和最整洁的可能的微分学的定义"是"在于它代数地确定了比值的极限,对此我们已经有了以直线形式的表达式,以及在于使那两个表达式相等."[46] 作为他所表示的意思的一个例子,达朗贝尔对于抛物线 $y^2=ax$ 通过首先确定经过两点 (x,y) 和 $(x+u,y+z)$ 的割线的斜率计算切线的斜率.这个斜率即,比值 $z:u$,容易看出等于 $a:2y+z$."这个比值总小于 $a:2y$;但 z 越小,比值将越大,并且由于可以选取 z 愿意多么小就多么小,所能产生的比值 $a:2y+z$ 按我们希望与比值 $a:2y$ 要有多接近就有多接近.结果,$a:2y$ 是比值 $a:2y+z$ 的极限."[47] 由此得出,$dy/dx=a/2y$.达朗贝尔的措词和麦克劳林的在实质上是相同的.但是,通过在他的《百科全书》的文章中给出"极限"术语的明确的定义,他走得稍远些.文章关于那个概念写到:"一个量被称作另一个量的**极限**是指当第二个可以在任意给定的量内接近第一个,不论多么小,第二个量绝不可以超过它所接近的量."[48] 他的思想,虽然明显是几何的而不是算术的,并没有被18世纪的他的继承人遵循.贯穿这个世纪剩余部分,大多数著作试图用无穷小量、流数或零的比值的方式解释这门学科的基础.

13.5.4　拉格朗日和幂级数

人物小传	约瑟夫·路易斯·拉格朗日(1736—1813)(Joseph-Louis Lagrange)
	拉格朗日(图 13.15) 出生于杜林的一个祖籍法国的家庭. 他的父亲希望他学习法律, 但是他被吸引到数学并且在 19 岁成为在杜林的皇家炮兵学校的一名数学教授. 在大约同一时期, 在已经读过欧拉的关于变分学的书籍, 他给欧拉写信解释他最近发现的得出这门学科的中心方程的一个更好的方法. 欧拉非常称赞拉格朗日并且准备把他的论文递交柏林科学院. 弗里德里克二世也对拉格朗日的工作印象深刻, 并且当欧拉离开他在科学院的职位回到圣彼得堡时, 弗里德里克任命拉格朗日继承这个职位. 在弗里德里克去世之后, 拉格朗日接受了路易十六的邀请来到巴黎, 在那里他度过他生命的其余时光, 1788 年在那里出版了他最重要的著作,《分析力学(Analytical Mechanics)》. 这部著作延续了牛顿、伯努利和欧拉的力学, 并且强调在力学中的问题通过简化它们到常微分和偏微分方程一般可以被解决的事实. 在 1792 年他与 17 岁的瑞纳·莫尼尔结婚, 她给他的生活带来重新的欢乐. 他有一种平常内向的性格, 这或许使他经过法国大革命的暴行仍然活下来. 他事实上被以荣誉相待, 但是他的几个同事的死亡极大地干扰了他. 在这次恐怖之后, 他在改进法国的大学教育中发挥积极作用并且最后被拿破仑为他一生的工作授勋.

拉格朗日在临近 18 世纪末试图通过消除所有涉及无穷小量、流数、零, 甚至极限的量, 他认为它们都缺少妥当的定义, 并以此给出导数的准确的定义. 他在 1772 年的一篇论文中勾画了他的关于导数的新思想, 然后在他的 1797 年的课本中全面发展了这些新思想. 课本的完整标题表达了他想要做什么: *解析函数的理论, 包括微分学的原理, 从无穷小量或渐近于零的量、极限或流数的每一种考虑中解脱出来, 简化到有限量的代数分析*. 拉格朗日如何能够完成把微积分学简化到纯代数的分析? 他通过把大多数前辈们没有疑问地使用的一个思想形式化来完成它, 这个思想就是任意函数都可以表示成幂级数. 对于拉格朗日而言, 若 $y = f(x)$ 是任一函数, 则 $f(x + i)$, 其中 i 是一个不确定的量, 可以"利用级数理论"展开成的级数:



图 13.15　法国邮票上的拉格朗日.

$$f(x + i) = f(x) + pi + qi^2 + ri^3 + \cdots,$$

其中 $p, q, r, \cdots$ 是关于 x 的不依赖于 i 的新函数. 拉格朗日然后证明比值 dy/dx 可以在这个展开中用 i 的一次幂的系数 $p(x)$ 来表示. 他因此有了微积分学的这个基本概念的一个新定义. 由于函数 $p(x)$ 是从原来的函数 f "得来"的, 拉格朗日称它是一个导出的函数(fonction dérivée)(由此出现英语词汇导数 derivative) 并使用符号 $f'(x)$, 类似地, f' 的导数写为 f'', f'' 的导数写为 f''', 等等. 拉格朗日容易证明 $q = (1/2)f''$, $r = (1/6)f'''$, $\cdots$.

为了回答拉格朗日为什么相信每一函数可以被展开成幂级数的这个明显的问题, 我们必须先考虑拉格朗日的函数的定义, 它出现于他的课本的最开始: "一个称为另一个或几个量的**函数**是指任意的数学表达式, 在其中这些量以任何方式进入, 无论是否与其他已给的或常数的值有关, 而函数的量可以取任何可能的值."[49] 换句话说, 拉格朗日本质上已经回到了欧拉对函数的第一个定义上, 剩下的是有点模糊的"数学表达式"和"以任意方式"的观念. 拉格朗日对函数的经验告诉他, 代数表达式总是可以展开成幂级数的: "这个假设实际上被通过不同的已知函数的展开所验证; 但是

据我所知,还没有人寻求去先验地证明它."[50] 拉格朗日认识到对某个特殊的 x 值,这样一个幂级数可能不存在,但对他而言,特殊的例外的值并不重要.他强调 x 以及 i 都被看作是不定的,使得 $f(x)$ 和 $f(x+i)$ 只是形式的表达式因而不能取值为无穷.

拉格朗日对函数 f 的展开式的论证从断言 $f(x+i)=f(x)+iP$ 开始,其中 $P(x+i)$ 定义为

$$P(x,i)=\frac{f(x+i)-f(x)}{i}.$$

拉格朗日进一步假设在 $i=0$ 处不等于零的那部分 p 可以从 P 中分离出来.亦即,$p(x)$ 定义为 $P(x,0)$,并且

$$Q(x,i)=\frac{P(x,i)-p(x)}{i},$$

或 $P=p+iQ$. 由此得出 $f(x+i)=f(x)+ip+i^2Q$. 对 Q 重复这个论证,他记 $Q=q+iR$,并再替代.作为这个过程的一个例子,拉格朗日取 $f(x)$ 为 $1/x$. 由于 $f(x+i)=1/(x+i)$,他计算出

$$P=\frac{1}{i}\left(\frac{1}{x+i}-\frac{1}{x}\right)=-\frac{1}{x(x+i)},\quad p=-\frac{1}{x^2},$$

$$Q=\frac{1}{i}\left(-\frac{1}{x(x+i)}+\frac{1}{x^2}\right)=\frac{1}{x^2(x+i)},\quad q=\frac{1}{x^3}.$$

$$\cdots$$

因此级数变成

$$\frac{1}{x+i}=\frac{1}{x}-\frac{i}{x^2}+\frac{i^2}{x^3}-\frac{i^3}{x^4}+\cdots.$$

在展开的每个阶段,这些项 $iP,i^2Q,\cdots$ 可以被认为是从用关于那个点的各项表示 $f(x+i)$ 产生的误差项.进而拉格朗日宣称,i 的值总可以取得这样小,使得这个级数中的任意给定项大于剩余项之和,亦即,余项总是充分小使得实际上函数是由级数表示的.事实上,拉格朗日后来经常使用这个结果.他还使用一个稍微不同的含有现在被称为拉格朗日型余项的泰勒级数的展开结果.亦即,他证明对于任意给定的正整数 n,可以写出

$$f(x+i)=f(x)+if'(x)+\frac{i^2f''(x)}{2}+\cdots+\frac{i^nf^{(n)}(x)}{n!}+\frac{i^{n+1}f^{(n+1)}(x+j)}{(n+1)!},$$

对于在 0 和 i 之间的某个 j 值成立.虽然这个新形式对于现代的读者来说或许并不比对于先前的读者来说更令人信服,拉格朗日自己满意于对于每个函数,他的幂级数表示的原理是正确的.最终他宣称,它使他能够不用考虑无穷小量、流数或极限重新得出微积分学的所有基本结果.

这些基本结果之一就是今天作为微积分学基本定理所知道的一部分,即若 $F(x)$ 表示从一个固定纵坐标在曲线 $y=f(x)$ 下方的面积,则 $F'(x)=f(x)$.(需要指出的是拉格朗日没有面积的定义.他简单假定在曲线 $y=f(x)$ 下方的面积是确定好的量.)拉格朗日的证明,使人想起麦克劳林的对幂级数的同样的结果的证明,即从指出 $F(x+i)-F(x)$ 表示在横坐标 x 和 $x+i$ 之间的那部分面积开始.按照欧拉的格言在关于分析学的课本中不应该包括图示,尽管如此拉格朗日还是写到,即使没有图形仍然容易使自己相信,若 $f(x)$ 是单调增加的,则 $if(x)<F(x+i)-F(x)<if(x+i)$;若是单调减少的,则不等式反过来(图 13.16).

现在展开 $f(x+i)$ 和 $F(x+i)$,拉格朗日求出

$$f(x+i)=f(x)+if'(x+j),$$

和

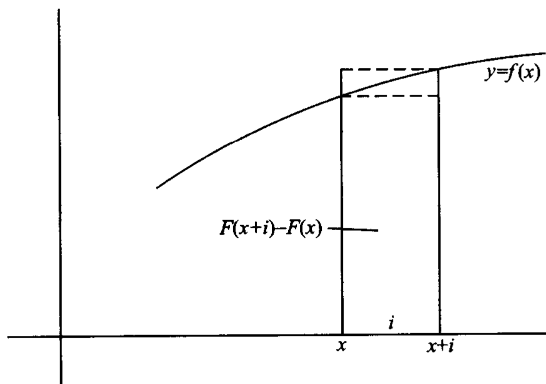


图 13.16 拉格朗日和微积分基本定理：$if(x) < F(x+i) - F(x) < if(x+i)$.

$$F(x+i) = F(x) + iF'(x) + \frac{i^2}{2}F''(x+j),$$

其中$0 < j < i$(虽然j的值在两个展开式中可以不是相同的). 由此得出 $if(x) < iF'(x) + \frac{i^2}{2}F''(x+j) < if(x) + i^2 f'(x+j)$,并且因此有

$$\left| i[F'(x) - f(x)] + \frac{i^2}{2}F''(x+j) \right| < i^2 f'(x+j),$$

其中为了照顾到增加和减少两种情形,绝对值的符号是必须的. 拉格朗日得出结论,因为这个不等式无论 i 取多么小都成立,则 $F'(x) = f(x)$ 必然正确. 他甚至求出,假若这个结论不正确,则因

$$i < \frac{F'(x) - f(x)}{f'(x+j) - \frac{1}{2}F''(x+j)},$$

而使不等式不成立. 为了完成他的证明,拉格朗日去掉了函数 $f(x)$ 在原来的区间$[x, x+i]$上单调的条件. 因为若它不是单调的,则在区间$[x, x+i]$上有 f 的一个最大值或者最小值,并且 i 可以被选得足够小使得极值落在新区间$[x, x+i]$之外.

这当然是奇怪的,尽管拉格朗日宣称这本著作将只用"有限量的代数分析,"他事实上在他的这个非常重要的证明和在他的余项的论证中已经运用了极限的观念. 和其他的几何量一样,在这本书的求切线、曲率、最大值和最小值的其他章节,他涉及到幂级数的函数的展开式的中心概念的同时也运用极限. 事实上,正是这些论证被用于 19 世纪的明确地运用极限的微积分学的研究中.

大多数早期的对拉格朗日的微积分学新基础的反对意见指向他的新符号和某些运算的长度,而不是指向他的任意函数都可以展开成幂级数的论断. 数学家们一般继续使用较早的微分方法,尤其是自从拉格朗日的书向他们保证因为整个学科有一个正确的基础,它做出的任何方法都是正当的. 甚至拉格朗日在他的某些其他著作中继续使用微分的概念而不是导数的概念. 直到 19 世纪的第二个十年才有各种不同的数学家指出存在可微的而没有幂级数表达式的函数因而拉格朗日的基本概念并不是站得住脚的. 给微积分学的思想提供基础的新尝试的故事因此将在 16 章中继续.

习　题

微分方程中的问题

注:在这一节中的问题一般需要微分方程学科的基本知识.

1. 从约翰·伯努利的微分方程 $dy = \frac{a dx}{\sqrt{x^2 - a^2}}$ 推导闭形式的悬链线方程.
2. 从雅各布·伯努利的方程 $ds = \frac{\sqrt{a} dy}{\sqrt{x}}$ 推导出对最速降线的微分方程 $dy = dx\sqrt{\frac{x}{a - x}}$.
3. 设已知的函数 f 是代数的,确定求已知族 $f(x, y, \alpha) = 0$ 的正交轨线族的微分方程一个方法.(利用正交线有负倒数斜率的事实.)使用你的方法求正交于双曲线族 $x^2 - y^2 = a^2$ 的曲线族.
4. 确定和求解等时线族,即正交于最速降线族的曲线族的微分方程.
5. 假设 $A = a_1 + a_2 + \cdots + a_n$ 和 $B = b_1 + b_2 + \cdots + b_n$. 证明 $\sum(b_i - a_i) = B - A$,或者说部分差的和等于部分和的差.
6. 将莱布尼茨的 $m dx + ny dx + dy = 0$ 的解答翻译成现代术语,注意 $dp/p = n dx$ 等于 $\ln p = \int n dx$ 或 $p = e^{\int n dx}$. 利用莱布尼茨的方法求解 $-3x dx + (1/x) y dx + dy = 0$.
7. 证明在齐次方程 $dy = f(y/x) dx$ 中可以通过利用代换 $y = vx$ 分离变量. 应用这一技巧求解方程 $x^2 dy = (y^2 + 2xy) dx$.
8. 证明 $y = e^{x/a}$ 是微分方程 $a d^3 y - y dx^3 = 0$ 的解. 然后,假设乘积 $e^{-(x/a)}(a d^3 y - y dx^3)$ 是 $e^{-(x/a)}(A d^2 y + B dy dx + C y dx^2)$ 的微分并且原方程的一个新解也一定满足 $a^2 d^2 y + a dy dx + y dx^2 = 0$.(提示:求微分并且使两式相等. 如果你利用导数以现代符号改写会更容易.)
9. 已知 $y = e^x$ 是 $y''' - 6y'' + 11y' - 6y = 0$ 的一个解,利用习题 8 的方法的类似方法证明任何其他解一定满足 $y'' - 5y' + 6y = 0$.
10. 通过利用欧拉的因式分解特征多项式的方法,求解习题 9 的微分方程.
11. 证明如果 $y = ue^{\alpha x}$ 被假定为 $a^2 d^2 y + a dy dx + y dx^2 = 0$ 的一个解,那么若 $\alpha = -(1/2)a$,则得出结论 u 是 $a^2 d^2 u + (3/4) u dx^2 = 0$ 的一个解.
12. 求解 $a^2 d^2 u + (3/4) u dx^2 = 0$. 先乘以 du,并且积分一次,得到 $4a^2 du^2 = (K^2 - 3u^2) dx^2$ 或 $dx = \frac{2a}{\sqrt{K^2 - 3u^2}} du$,第二次积分得到 $x = \frac{2a}{\sqrt{3}} \arcsin \frac{\sqrt{3} u}{K} - f$. 改写这个方程为 u 用 x 来表示如下:
$$u = K \sin\left(\frac{(x + f)\sqrt{3}}{du}\right).$$
13. 因为任意复数 z 可以写作 $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$ 其中 $r > 0$,可以定义 $\log z = \log re^{i\theta}$ 为 $\log r + i\theta$, 其中 $\log r$ 是正数 r 的实对数. 证明因此一个复数有无穷多对数. 在什么情形下复数的任何对数是实数?
14. 从 13.12 节利用克莱罗的方法求解微分方程 $(2xy^3 + 6x^2y^2 + 8x) dx + (3x^2y^2 + 4x^3y + 3) dy = 0$.
15. 求在上半平面连接两点的曲线,当绕 x 轴旋转时,生成最小表面积的曲面. 如果 $y = f(x)$ 是曲线的方程,那么所求表面积为 $I = 2\pi\int y ds = 2\pi\int y\sqrt{1 + y'^2} dx$. 所以可以利用修改形式的欧拉方程 $F - y'(\partial F/\partial y') = c$,其中 $F = y\sqrt{1 + y'^2}$.(提示:从方程两边乘以 $\sqrt{1 + y'^2}$ 开始.)

从微积分课本中来的问题

16. 求外接半径为 1 的圆的面积最小的等腰三角形.(辛普森)
17. 求具有已知体积的表面积最小的圆锥.(辛普森)

18. 证明 $w = (b^3 - x^3)(x^2z - z^3)(xy - y^2)$ 当 $x = \frac{1}{2}b\sqrt[3]{5}, y = \frac{1}{4}b^3\sqrt{5}$ 和 $z = \frac{b\sqrt[3]{5}}{2\sqrt{5}}$ 时有最大值.(辛普森)
19. 利用麦克劳林的技巧而不明显地利用正弦和与余弦函数的导数来计算 $y = \cos z$ 的幂级数的前四个非零项.假定圆半径为1.
20. 在微分方程 $a(\mathrm{d}y/y) = \mathrm{d}x$ 定义的曲线上求曲率最大的点.(艾格纳丝)
21. 已知一个矩形,求经过一个角点以最短长度穿过两条对边的延长线的那条直线(艾格纳丝).
22. 画出以 $y^2 = \frac{4(2-x)}{x}$ 给出的"艾格纳丝巫婆"的一个特殊的例子.证明它对称于 x 轴并且渐近于 y 轴.证明在曲线和其渐近线之间的面积是 4π.
23. 假设在大洪水后人口为6并且在200年以后人口为1 000 000.求人口的年增长率(欧拉).提示:如果年增长率是 $1/x$,那么问题的方程是

$$6\left(\frac{1+x}{x}\right)^{200} = 1\,000\,000.$$

24. 证明在正文中给出的在 a^x 和 $\log_a(1+x)$ 的幂级数中的 k 可以用 $k = \ln a$ 表示.(欧拉)
25. 从方程 $\log(1+x) = j(1+x)^{1/j} - j$ 利用二项式定理并假设 j 是无穷大推导 $\ln(1+x)$ 的幂级数.(欧拉)
26. 如果 $y = \arctan x$,证明 $\sin y = x/\sqrt{1+x^2}$ 和 $\cos y = 1/\sqrt{1+x^2}$.然后,如果 $p = x/\sqrt{1+x^2}$,证明 $\sqrt{1-p^2} = 1/\sqrt{1+x^2}$.因为 $y = \arcsin p$,由此得到 $\mathrm{d}y = \mathrm{d}p/\sqrt{1-p^2}$ 和 $\mathrm{d}p = \mathrm{d}x/(1+x^2)^{3/2}$.得出结论 $\mathrm{d}y = \frac{\mathrm{d}x}{1+x^2}$.(欧拉)
27. 确定 $V = x^3 + y^2 - 3xy + (3/2)x$ 的所有的相对极值,并且对于每个极值确定它是否是极大值或极小值.将你的答案和欧拉的相比较.
28. 对于 $y = a^x$,注意到 $\mathrm{d}y = a^{x+\mathrm{d}x} - a^x = a^x(a^{\mathrm{d}x} - 1)$,并且然后展开 $a^{\mathrm{d}x} - 1$ 成为幂级数 $\ln a\mathrm{d}x + \frac{(\ln a)^2\mathrm{d}x^2}{2} + \cdots$,计算 $\mathrm{d}y$.(欧拉)
29. 对于 $y = \tan x$,利用加法公式 $\tan(x + \mathrm{d}x) = \frac{\tan x + \tan \mathrm{d}x}{1 - \tan x\tan \mathrm{d}x}$,计算 $\mathrm{d}y$.(欧拉)
30. 已知 $z^n - a^n$ 的根是 $a\left(\cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n}\right), k = 0,1,\cdots,n-1$,并且复根总是成对出现,证明当 $1 \leqslant k \leqslant \frac{n}{2}$ 时 $z^2 - 2az\cos\frac{2k\pi}{n} + a^2$ 是 $z^n - a^n$ 的实二次因式.(这个习题和以下四个习题提供欧拉的整数平方的倒数和的值的推导.)
31. 回顾当 ω 无穷小和无穷大时 $a^{j\omega} = (1 + k\omega)^j$.令 $j\omega = x$(有限值),$a = e$ 和 $k = 1$.则 $\mathrm{e}^x = \left(1 + \frac{x}{j}\right)^j$ 和类似地 $\mathrm{e}^{-x} = \left(1 - \frac{x}{j}\right)^j$ 由习题30证明 $\mathrm{e}^x - \mathrm{e}^{-x}$ 和当 $k = 1,2,\cdots$ 时一切形式 $\frac{4x^2}{j^2} + \frac{4k^2\pi^2}{j^2} - \frac{4k^2\pi^2x^2}{j^4}$ 的二次因式.(提示:因为 j 是无穷大,仅利用 $\cos\frac{2k\pi}{j}$ 式的前两项.)
32. 由习题30证明,因为在三项式因式中第三项相对于其他两项是无穷小,当 $k = 1,2,\cdots$ 时 $\frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}$ 是用 $1 + \frac{x^2}{k^2\pi^2}$ 可除尽的.得出结论

$$\frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2} = x\left(1 + \frac{x^2}{\pi^2}\right)\left(1 + \frac{x^2}{4\pi^2}\right)\left(1 + \frac{x^2}{9\pi^2}\right)\cdots.$$

33. 在习题31中以 $\mathrm{i}x$ 代替 x 得出结论

$$\sin x = \frac{\mathrm{e}^{\mathrm{i}x} - \mathrm{e}^{-\mathrm{i}x}}{2\mathrm{i}} = x\left(1 - \frac{x^2}{\pi^2}\right)\left(1 - \frac{x^2}{4\pi^2}\right)\left(1 - \frac{x^2}{9\pi^2}\right)\cdots,$$

注意到这个乘积表达式证明 $\sin x$ 在 $0, \pm\pi, \pm 2\pi, \pm 3\pi, \cdots$ 有根.
34. 因为

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots = x\left(1 - \frac{x^2}{6} + \frac{x^4}{120} - \cdots\right),$$

从在习题 33 的乘积表达式得出结论

$$\sum_{k=1}^{\infty}\frac{1}{k^2\pi^2}=\frac{1}{6} \quad 或 \quad \sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$$

这个结果,被雅各布和约翰·伯努利寻求,在 1753 年由欧拉第一次证明.

二重积分

35. 假设 x 和 y 由函数 $x=\dfrac{t}{\sqrt{1+u^2}}, y=\dfrac{tu}{\sqrt{1+u^2}}$ 用 t 和 u 的形式表出.证明变量变换公式表示为

$$\mathrm{d}x\mathrm{d}y=\frac{t\mathrm{d}t\mathrm{d}u}{1+u^2}.$$

36. 利用克莱罗技巧计算由柱面 $ax=y^2$、$by=z^2$、坐标面和平面 $x=x_0$ 围成的立体的体积.首先,通过将被积表达式变换成为 x 的函数并且进行积分,确定体积元素 $\mathrm{d}x\int z\mathrm{d}y$.然后,用适当的限对体积元素积分.将这个方法与标准的现代方法相比较.

波动方程

37. 假设波动方程 $\dfrac{\partial^2 y}{\partial x^2}=\dfrac{\partial^2 y}{\partial x^2}$ 的解被给定为 $y=\Psi(t+x)-\Psi(t-x)$.证明初始条件 $y(0,x)=f(x)$, $\dfrac{\partial y}{\partial t}(0,x)=g(x)$ 和对于一切 t, $y(t,0)=y(t,l)=0$ 的条件导至要求 $f(x)$ 和 $g(x)$ 是周期为 $2l$ 的奇函数.(达朗贝尔)

38. 假设 $y=F(t)G(x)=\Psi(t+x)-\Psi(t-x)$ 是波动方程 $\dfrac{\partial^2 y}{\partial t^2}=\dfrac{\partial^2 y}{\partial x^2}$ 的一个解.通过两次微分证明 $\dfrac{F''}{F}=\dfrac{G''}{G}=C$,其中 C 是某个常数,由此有 $F=ce^{t\sqrt{C}}+de^{-t\sqrt{C}}$ 和 $G=c'e^{x\sqrt{C}}+d'e^{-x\sqrt{C}}$.应用条件 $y(t,0)=y(t,l)=0$,证明 C 一定是负的,以及因此对于适当选取的 A、B 和 N 得出解 $F(t)=A\cos Nt$ 和 $G(x)=B\sin Nx$.(达朗贝尔)

微积分学的基础

39. 翻译达朗贝尔的极限定义成为代数语言并与现代的极限定义相比较.
40. 使用拉格朗日的技巧计算函数 $f(x)=\sqrt{x}$ 的量 p、q 和 r,由此确定它的幂级数表达式的前三项.
41. 说明为什么幂级数表达式对于 $f(x)=\mathrm{e}^{-1/x^2}$ 的情形失效.
42. 已知 $f(x+i)=f(x)+pi+qi^2+r^3+\cdots$,证明 $p=f'(x)$, $q=f''(x)/2!$, $r=f'''(x)/3!$, $\cdots$.

讨 论

43. 是否 18 世纪的数学家在今天使用微积分基本定理的意义下证明或利用它?甚至只在考虑这个定理之前就必须定义一些什么概念?18 世纪的数学家是如何处理这些概念的?是否这些数学家认为这个基本定理是"基本的"?
44. 为了增进学生对于微积分基本定理的理解,通过利用麦克劳林和拉格朗日的工作设计一堂课.
45. 为了教授在重积分中的变量替换定理,利用欧拉技巧设计一堂课.
46. 将欧拉的预微积分学和微积分学课本的三部曲与现代系列相比较.共同的是什么项目?在今天的课本里没有的而欧拉有的是什么,反过来呢?在今天有人能用欧拉的课本吗?
47. 追录从牛顿,经麦克劳林,一直到达朗贝尔的关于极限概念的发展路经.它们的公式化表述是如何一致的?与此概念现代表述比较起来它们又如何?
48. 制定多个课来讲授用莱布尼茨,克莱罗和欧拉的公式化表述去解决各类微分方程的基本方法.

文献和注解

虽然没有详细地处理本章的所有主题的单一的著作,但数本书处理此内容的某些方面.它们包括 S. B. Engelsman, Families of Curves and the Origins of Partial Differentiation(Amsterdam: Elsevier, 1984), Umberto Bottazzini, The Higher Calculus: A History of Real and Complex Analysis from Euler to Weierstrass(New York: Springer-Verlag, 1986),以及 Ivor Grattan-Guinness, The Development of the Foundations of Mathematical Analysis from Euler to Riemann(Cambridge: MIT Press, 1970). 两篇好的处理此内容的概览文章是 H.J.M.Bos, "Calculus in the Eighteenth Century-The Role of Applications,"Bulletin of the Institute of Mathematics and Its Applications 13(1977), 221 – 227 和 Craig Fraser, "The Calculus as Algebraic Analysis: Some Observations on Mathematical Analysis in the 18th Century,"Archive for History of Exact Sciences 39(1989), 317 – 336.

1. David Eugene Smith, *A Source Book in Mathematics*(New York: Dover, 1959), vol.2, p.646.

2. 同上, p.645. 原问题和约翰·伯努利的解在这里给出.

3. 同上, p.647.

4. 同上, p.654.

5. 雅各布·伯努利对最速降线问题的解出现在 D.J.Struik, *A source Book in Mathematics*, 1200—1800(Cambridge: Harvard University Press, 1969), pp.396 – 399.

6. Engelsman, *Families of Curves*, p.106. 这本著作奉献最好的偏微分法的历史和这一节的许多内容的来源.

7. C.Truesdell, "The Rational Mechanics of Flexible or Elastic Bodies: 1638—1788," 在 Leonhard Euler, *Opera Omnia*(Leipzig, Berlin, and Zurich: Societas Scientarum Naturalium Helveticae, 1911—), (2)11, part 2, p.166. 这篇文章包括关于欧拉的工作的许多细节.它是欧拉文集一部分的介绍,现在总共超过 70 卷,耗时 80 年以上用四个系列出版并仍未完成,所有的欧拉的论文的参考文献在这个集子中将含有它们的出处.关于正弦和余弦函数的微积分学的发明的更详尽的研究可见 Victor J. Katz, "The Calculusof the Trigonometric Functions,"Historia Mathematica14 (1987), 311 – 324. 关于欧拉的工作的不同的文章可在他逝世两周年的特刊中找到: *Mathematics Magazine* 56 (5) (1983). 还见 E.A.Fellmann, ed., Leonhard Euler 1707—1783: *Beiträge zu Leben und Werk*(Boston: Birkhäuser, 1983).

8. G.Eneström, "Der Briefwechsel zwischen Leonhard Euler und Johann Bernoulli," *Bibliotheca Mathematica* (3) 6 (1905), 16 – 87, p.13.

9. 同上, p.46.

10. John Fauvel and Jeremy Gray, *The History of Mathematics*: *A reader*(London: Macmillan, 1987), p.452. 这个引用是一个部分原文译文,它来自欧拉的, "De la Controverse entre Mrs Leibniz et Bernoulli sur les Logarithmes des Nombres Negatifs et Imaginaires," *Mem. Acad. Sci. Berlin*(1749) = Opera Omnia(1)vol.17, 195 – 232.

11. 欧拉的变分学的观点的这一研究归功于 American University 的 Steven Schot, 在某些未发表的讲义中.变分学的详尽的历史可以找 Herman H.Goldstine, *A History of the Calculus of Variations from the 17th through the 19th Century*(New York: Springer Verlag, 1980).

12. Frances Marguerite Clarke, *Thomas Simpson and his Times*(New York: Columbia University, 1929), p.16. 这本著作是惟一的托马斯·辛普森的传记并且提供他的生活和工作的许多细节.

13. Thomas Simpson, *A New Treatise of Fluxions*(London: Gardner, 1737), p.179.

14. Colin Maclaurin, *A Treatise of Fluxions*(Edinburgh: Ruddimans, 1742), p.694.

15. 同上, p.753. 关于麦克劳林的微积分学及其影响的更详细的情况,见 Judith Grabiner, "Was Newton's Calculus a Dead End? The Continental Influence of Maclaurin's Treatise of Fluxions," *American Mathematical Monthly* 104(1997), 393 – 410.

16. Edna Kramer, "Maria Agnesi," *Dictionary of Scientific Biograghy*, vol.1, p.76. 关于艾格纳斯的更近期的工作是 C.

Truesdell, "Maria Gaetana Agnesi," *Archive for History of Exact Sciences* 40(1989), 113 - 147.

17. Euler, *Introduction to Analysis of the Infinite, Book I*, John D. Blanton 译(New York: Springer Verlag, 1988), p.v. 这一工作是欧拉的《导论》的近期英文翻译的第一卷.(第二卷见注记26.)仔细阅读会有良好回报，因为在那里有很多课本里不能概括的东西. 还见 Car Boyer, "The Foremost Textbook of Modern Times(欧拉的 *Introductio in analysin infinitorum*)," *American Mathematical Monthly* 58(1951), 223 - 226.
18. 同上, p.3. 关于函数记号的更详细的情况，见 A.P. Youshkevitch, "The Concept of Function up to the Middle of the Nineteenth Century," *Archive for History of Exact Sciences* 16 (1982), 37 - 85.
19. Euler, *Institutiones Calculi Differentialis* 在 Opera Omnia(1) vol.10, p.5.
20. 同上, p.4.
21. 同上, p.84.
22. 同上, p.99.
23. D.J. Struik, *Source Book*, p.355. Struik 提供了有关弦振动争论的文章的重要部分的翻译. 全部事件的一个讨论可以找 Jerome R. Ravetz, "Vibrating Strings and Arbitrary Functions," 这一章见 *The Logic of Personal Knowledge: Essays Presented to Michael Polanyi on his Seventieth Birthdays*, 11 *March* 1961(London: Routledge and Paul, 1961), 71 - 88, 以及 Bottazzini, Higher Calculus 的第一章和 C. Truesdell, "Rational Mechanics." d'Alembert 的工作的全面的研究是 Thomas Hankins, *Jean d'Alembert: Science and the Enlightenment* (Oxford: Clarendon Press, 1970).
24. 引用 Truesdell, "Rational Mechanics," p.239.
25. 引用 Struik, *Source Book*, p.361.
26. Euler, *Introduction to Analysis of the Infinite, Book II*, John D. Blanton 译(New York: Springer Verlag, 1990), p.6.
27. 引用 Bottazzini, *Higher Calculus*, p.27.
28. 引用 Struik, *Source Book*, p.361.
29. George Berkeley, *The Analyst*, in James Newman, *The World of Mathematics* (New York: Simon and Schuster, 1956), vol.1, 288 - 289. 伯克利的著作的节选在此重印并且还见 Struik, Source Book, p.333 - 338.
30. 同上, pp.291 - 292.
31. Struik, Source Book, p.335.
32. Newman, *The World of Mathematics*, p.289.
33. 同上, p.290.
34. Maclaurin, *Treatise of Fluxions*, p.1. 柯林·麦克劳林的简短的传记见 H.W. Turnbull, *Bicentenary of the Death of Colin Maclaurin* (Aberdeen, University Press, 1951). 又见 C. Tweedie, "A Study of the Life and Writtings of Colin Maclaurin," *Mathematical Gazette* 8(1915), 132 - 151 和 9(1916), 以及 H.W. Turnbull, "Colin Maclaurin," *American Mathematical Monthly* 54(1947), 318 - 322.
35. 同上, p.4.
36. 同上, p.53.
37. 同上, p.55.
38. 同上, p.181.
39. 同上, p.420.
40. 同上, p.421.
41. 同上, p.423.
42. 同上, p.38.
43. Struik, *Source Book*, p.384. Struik 有一个从欧拉的《微分学》中的几页关于微积分的哲学讨论的翻译.
44. 同上.
45. 同上, p.385.
46. 同上, p.345. 这个节选来自于 d'Alembert 在 Encyclopédie 中的文章"Différentiel".
47. 同上, pp.343 - 344.

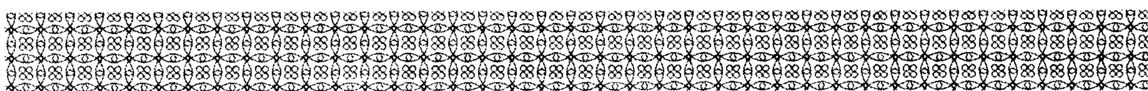
48. 来自于在《百科全书》中的文章“Limite.”

49. Lagrange, *Theory of Analytic Functions*, *in Oeuvres de Lagrange*(Paris:Gauthier-Villars,1881),vol.9,p.15.一个简短的节选翻译在 Struik, *Source Book*, pp.388 – 391.

50. 同上,p.22.

18 世纪分析概览

1622—1703	文森佐 · 维维安尼(Vincenzo Viviani)	重积分问题
1646—1716	奇特弗里德 · 魏尔海姆 · 莱布尼茨(Gottfried Wilhelm Leibniz)	偏导数、微分方程
1654—1705	雅各布 · 伯努利(Jakob Bernoulli)	最速降线问题
1667—1748	约翰 · 伯努利(Johann Bernoulli)	最速降线问题、微分方程
1685—1731	布鲁克 · 泰勒(Brook Taylor)	Taylor 级数
1685—1753	乔治 · 伯克利(George Berkeley)	微积分学基础的批评
1687—1759	尼古拉斯 · 伯努利(Nicolaus I Bernoulli)	偏导数法则
1698—1746	柯林 · 麦克劳林(Colin Maclaurin)	微积分学课本
1700—1782	丹尼尔 · 伯努利(Daniel Bernoulli)	弦振动问题
1707—1783	列昂哈德 · 欧拉(Leonhard Euler)	微分方程、教科书
1710—1761	托马斯 · 辛普森(Thomas Simpson)	微积分学课本
1713—1765	阿列克塞 · 克劳德 · 克莱罗(Alexis Claude Clairaut)	微分方程
1717—1783	让 · 隆德 · 达朗贝尔(Jean Le Rond d'Alembert)	弦振动问题
1718—1799	玛利亚 · 艾格纳丝(Maria Agnesi)	微积分学课本
1736—1813	约瑟夫 · 路易斯 · 拉格朗日(Joseph-Louis Lagrange)	幂级数形式的微积分学



第14章 18世纪的概率、代数和几何

> 无论怎样,想要就某事的发展做出一个正确的预测,似乎只需要严密地推算其各种可能出现的结果,然后确定其中一种比另一种更易发生的可能.然而,我们很快又会遇到麻烦,因为这种方法只在很少的一些情形下适用…….试问:谁人能够告知所有疾病的种类,推算出所有即将发生的病例……并说出一种病症比另一种更加致命的可能.或者,谁又能够列举出大气每日所经历的无数变化,并据以在当日预报出此后一个月,甚至一年内的天气?
>
> ——雅各布·伯努利:《猜度术》(1713)[1]

法国大革命结束以后,为了重建高等教育的需要,法国国家议会于1794年9月28日通过一项法案,决定创办巴黎综合工科学校(École Polytechnique).此后数十年间,法国所有最好的数学家们皆执教于此,他们中的大部分人还在此编写了使用的课本.综合工科学校很快就成为了整个欧洲以及美国的工程类大学的楷模.

尽管分析学的发展及其在众多领域内的应用形成了18世纪数学发展史上的核心内容,但其他领域也不乏重要成果.雅各布·伯努利继承了惠更斯在概率论方面的工作并对其进行了拓展.他给出了计算整数列幂和的一种新方法,还决定性地证明了现在所谓的大数定律.棣莫弗(Abraham De Moivre)借助于自己的级数知识,并通过从根本上解决正态分布曲线及其性质问题,将这一研究工作推向了更深层次.最终,贝叶斯(Thomas Bayes)和拉普拉斯证明了如何利用某些经验公式以确定概率.

这个世纪里还有几部主要的代数原著出版,牛顿、麦克劳林和欧拉的著作即在其中.这些著作对早期的研究成果进行了系统化总结.如麦克劳林的著作给出了解线性方程组的一种新方法,即通常所说的克拉默法则,而欧拉的著作则给出了数论中各种方法的一些具体结果.然而,代数学的一个主要目标仍在于将卡尔达诺和费拉里关于解方程的技术拓展到五次及更高次多项式方程.没有人在此方面有所成就.直至这个世纪末,拉格朗日详细考察了解三次和四次方程的方法,并提出一些自信在处理高次方程时必不可少的思想.

几何学中的两大推进也引人瞩目.首先是,基于与伊斯兰数学家们相似的工作,萨凯里和朗贝特发起了重新证明欧几里得平行公设的尝试.其次,更为重要的是,分析学开始应用于几何问题的

研究. 特别是,欧拉、克莱罗和蒙日对于解析几何和微分几何贡献了许多新思想. 欧拉还研究了两个特殊问题,其解答提供了下个世纪末拓扑学得以发展的思想萌芽.

正如开篇所言,法国大革命作为 18 世纪最为重要的政治事件,对于数学及其教育产生了重大影响. 因而本章将在结束前对综合工科学校开设的数学新课程作以简单回顾. 我们还将简要述及在新大陆(指美洲大陆)上数学发展的开端.

14.1 概率论

第 11 章论及的概率论早期著作所涉及的主要是针对于各种游戏和赌博问题而产生的期望值及其相关概率计算. 但是,帕斯卡关于游戏被中断后"公平"分配赌金的思想,以及惠更斯对"公平"赌博的兴趣,皆显示了概率论在其诞生之初与**投机性**合同概念的密切联系. 这种合同约定以现时某定值兑换未来某不确定值,如在年金保险和海洋运输保险的合同条款中,需在现时预付一定数量的保险金,才能在未来某种条件下获得一笔未知数量的回报. 为了使此类合同保持"公平",数学家们强调,人们需要有能力使用某种方法来量化所涉及的风险.

补遗 14.1

概率与现实世界

截止 18 世纪中叶,数学家们已经在概率论方面揭示出了足够多的结果,以致于人们或许会以为这些结果可以被应用于与实际收益相关的各种领域,如养老保险的定价或彩票奖额的设立. 然而令人惊讶的是,事实并非如此. 养老保险的发售史已有几个世纪,但它仍然被投保人普遍视作"压宝",而被售保人视为贷款. 也就是说,投保人付出固定数量的保金以换取终生获得定期报酬的保证,他们实际上赌的是自己能够活足够长的时间以收回所有报酬甚至获得更多. 另一方面,售保人则认为起初的保金不过是贷款,而自己付出的报酬就是利息,而且在通常情况下,这种利率要高于法定的贷款利率. 在死亡率统计表问世以前,为养老保险定价毫无数学方法. 因此,险金数额往往因有关当事人的经验或售保人对现金的需要来确定,而非出于统计资料. 即使在像威特和棣莫弗这样的数学家们基于对死亡率的统计给出养老保险表之后,惯例也未因之而发生变化. 例如,在英国,无论投保人的年龄如何,养老保险皆以每年回报的七倍数额被发售着. 而棣莫弗的统计表在此方面已经有所说明,如果利率被设定为5%,保险金的设定就应该在如下区间内浮动,即20岁的人应付年回报的13.89倍,而70岁的人则应付年回报的5.77倍. 另外,荷兰的一座小城镇曾通过以每份 250 荷兰盾的价格,售卖 400 份养老保险,从而筹集得 100 000 荷兰盾,然后尽量为那些依然存活着的投保人分配总共 4000 荷兰盾的年回报.

类似地,尽管数个世纪以来,各国政府一直通过设立摸彩的方式筹集资金,而且这一点在 17 世纪末和 18 世纪初变得相当流行,但这些政府在确定奖金数额时似乎并没有进行概率计算. 对于这些政府而言,通常似乎没有理由需要这样来做. 彩票业是政府收益的主要来源,因为考虑到中彩的机率,奖金并不算多. 例如,在法国皇家彩票的抽奖过程中,一位从 90 个数码中按照应有的顺序抽得 5 个正确数码的人,将获得彩票面值1 000 000倍的奖金. 但他赢得此奖的机率却远低于五十亿分之一. 显然,法国人尤其是那些社会下层人士,就像如今的低收入者,都愿意为此极其微小的变富机率而赌上几回. 虽然数学家们会不时地发表一些学术论文以告诉人们为什么不应该参与此类摸彩游戏,但人们只是对那些兜售如何才能确保抽得中奖号码的数学家们感兴趣.

在某种类型的赌博游戏中,早期的参与者们已能找出有效的途径来计算胜负,从而先验地推算出期望或概率. 然而在现实生活中,量化风险却并非易事. 也就是说,很难确定一位"理智的人"所

具有的自信程度.一个人应该如何来为支付保险而确定其"合理"的价钱呢?(参看补遗 14.1)正如开篇引语所言,雅各布·伯努利在对这一课题长达20多年的研究中,试图能够在无法列举出所有可能性的情形下量化风险.为了达到这一目的,他提议着眼于在众多相同情形下观察所得的结果来后验地计算概率."例如,倘若我们观察到在与某位提提乌斯具有相同体质的300位同龄人中,200人会在十年内死去,而其他人则得以继续生存,那么,我们就有理由认为,这些具有提提乌斯之体质的人在随后的十年间死亡的机率是其活过这十年的机率的两倍.[2]

14.1.1 雅各布·伯努利及其《猜度术》

越能更多地观察特定的情形,就越能更好地预测未来,这对于伯努利而言似乎是相当明显的事实.但他仍期望就此原理给出一个"科学地证明",从而不仅能说明观察次数的增加可以使得在任何精度范围内估算事件的实际概率成为可能,而且还可以说明如何准确地计算出必需进行多少次观察,才能确保估算结果位于真值预先给定的邻域内.于1705年他临去世前,伯努利在他的大数定律中已给出了这一科学证明,其方法被收入了他直到1713年才公开出版的重要的概率论著作《猜度术》(Ars Conjectandi).

大数定律出现在《猜度术》的第四卷也是最后一卷.前三卷论述的主要是一些早期概率论思想.事实上,第一卷在本质上相当于惠更斯在1657年的工作的再版,只是附加了一些注释说明;第二卷重新论述了各种组合学定律,它们中的大多数在数世纪以前已为人所知;而第三卷仅是借用这些定律去解决更多的赌博问题.然而在伯努利的著作中仍有两点原创性的工作值得一提.首先,他推广了帕斯卡关于游戏中断后如何分配赌金的思想.帕斯卡认为参与游戏的赌徒赢得所给分值的机会是均等的,而伯努利则将此推广到了赌徒们机会不等的情形,或者更为一般地,他的思想适用于胜负机会不等的任何试验.伯努利证明了,如果获胜的概率是 a,而失败的概率是 b(在 $a+b$ 次试验中),那么在 n 次试验中有 r 次获胜的概率为 $\binom{n}{n-r}a^rb^{n-r}\Big/(a+b)^n$.类似地,在 n 次试验中至少有 r 次获胜的概率为 $\sum_{j=0}^{n-r}\binom{n}{j}a^{n-j}b^j\Big/(a+b)^n$.

伯努利的第二个创新虽然用到了帕斯卡的算术三角形,但给出的却是计算整数列幂和的新方法.他在此方面超越了海塞姆和杰斯特德维:不仅得到了整数列直到10次的幂和公式,而且针对任何幂次 c 给出一种获取一般结果的通式:

$$\sum_{j=1}^{n}j^c=\frac{1}{c+1}n^{c+1}+\frac{1}{2}n^c+\frac{c}{2}B_2n^{c-1}+\frac{c(c-1)(c-2)}{2\cdot3\cdot4}B_4n^{c-3}$$
$$+\frac{c(c-1)(c-2)(c-3)(c-4)}{2\cdot3\cdot4\cdot5\cdot6}B_6n^{c-5}+\cdots,$$

其中右式以 n 的最后一个正数次幂为结尾,而 $B_2=1/6,B_4=-1/30,B_6=1/42,\cdots$ 这些如今被称作伯努利数的量,或许可以如此计算:注意到一旦求和时出现了某个伯努利数,那么该数与其前边 n 的幂次的各项系数总和应该为1.因而,由于

$$\sum j^4=\frac{1}{5}n^5+\frac{1}{2}n^4+\frac{1}{3}n^3+B_4n,$$

所以 $B_4=1-\frac{1}{5}-\frac{1}{2}-\frac{1}{3}=-\frac{1}{30}$.[3]

《猜度术》的第四卷题为"论概率原则的政治、伦理和经济学应用"(On the Use and Applications of the Doctrine in politics, Ethics, and Economics).虽然伯努利在事实上并没有讨论任何实践应用,

但他的确讨论了在现实生活中所观察到的各种迹象,以及这些迹象如何被组织成一条单纯的可能性陈述.意识到在大多数现实条件下,绝对确定(或者说概率为 1)是不可能实现的,伯努利引入了近乎必然性(moral certainty)这一概念.他规定对于一个近乎必然的结果,其出现的概率不应该小于0.999.相反地,如果某结果出现的概率不大于0.001,那它就是近乎不可能发生的.正是出于确定事件真实概率的近乎必然性的目的,伯努利导出了自己的定理,即大数定律.

为了理解定理的论述,我们必须时刻牢记伯努利给出的例题之一.设想一个盒子装有 3000 块白色小鹅卵石和 2000 块黑色小鹅卵石,但观察者对此数目并不知晓.观察者为了确定盒中黑白石子的比例,而按次序、有放回地抽取一定数量的小鹅卵石并记录下观察结果.如此,在下文中,每次观察事件为抽取一个石子,如果石子是白色的即算事件获得成功.现在我们假设,在一般情况下,N 次观察中有 X 次成功,而 $p = r/(r+s)$ 是(未知的)成功率.(这里的 r 代表所有的成功事件,s 代表所有的不成功事件,例题中的 $p = 3/5$).伯努利的定理用现代术语表述即是,给定一个任意小的分数 ε(伯努利通常愿意用 $1/(r+s)$)和一个任意大的正数 c,总能找到整数 $N = N(c)$,使得 X/N 与 p 的差距不超过 ε 的概率比该差距大于 ε 的概率乘以 c 还要大.这一结论可用符号表示为

$$p\left(\left|\frac{X}{N}-p\right|\leqslant\varepsilon\right)>cP\left(\left|\frac{X}{N}-p\right|>\varepsilon\right).$$

换句话说,X/N 接近于 p 的概率远远大于不接近的概率.在现代文献中,这一定理常常被表述为:任给 $\varepsilon>0$,c 为任意正数,存在正整数 N 使得

$$P\left(\left|\frac{X}{N}-p\right|>\varepsilon\right)<\frac{1}{c+1}.$$

由于概率的计算涉及到了二项展开式 $(r+s)^N$ 中某些项的和,伯努利详细分析了展开式中的所有项.这不仅使他找到了一个关于该定理的证明,而且给了他一种确定 $N(c)$ 的方法.特别地,他证明了如果 $t = r+s$,那么 $N(c)$ 可以取大于

$$mt+\frac{st(m-1)}{r+1}\quad 和\quad nt+\frac{rt(n-1)}{s+1}$$

两者中较大者的任何整数.其中 m,n 是整数,且满足如下条件

$$m\geqslant\frac{\log c(s-1)}{\log(r+1)-\log r}\quad 和\quad n\geqslant\frac{\log c(r-1)}{\log(s+1)-\log s}.$$

在自己的例题中,伯努利计算出,当 $r=30$,$s=20$ 时,第二个表达式更大,因而对于 $c=1000$,可以取 $N = 25\,550$.换句话说,伯努利的计算结果使得他能够确定,从 25 550 次观察中足以获知,观察所得成功率与真实概率 3/5 的误差不超过 1/50,该事件的发生具有近乎必然性.伯努利著作的结尾处除了该计算外,还有另一个针对不同 c 值的类似计算,这或许是因为他对于自己的计算结果并不满意.其原因在于 25 550 对于 18 世纪早期的人们而言是一个过于庞大的数字,例如,它甚至大于巴塞尔的人口数字.计算结果似乎在告诉人们,通过次数合乎情理的实验得不到值得信赖的结论.伯努利或许感觉到自己在寻求不确定量的量化测定过程中并不成功,尤其是他的直觉告诉自己,25 550 远大于必需的次数.[4] 因而他并没有把已经允诺的关于其方法在政治、经济学上的应用写入书中.[5] 不过,伯努利指明了一条道路,据此向这些问题发起成功进攻的,是一位比他稍微年轻的同代人,亚伯拉罕 · 棣莫弗(Abraham De Moivre,1667—1754).

14.1.2　棣莫弗及其《机会学说》

棣莫弗的主要数学著作《机会学说》首次公开出版于 1718 年,在 1738 年和 1756 年曾两次再版.

这一概率论著作远比惠更斯的著作更详细，部分原因在于1657年以来数学所发生的普遍进展.棣莫弗在给出一般法则之外，还通常就在他那个时代广泛流行的各种赌博游戏的玩法，给出这些法则的详细应用.例如德梅雷的掷骰子问题就曾被他视作某些更普遍的问题的特例而获得解决.

问题III 假设某事件发生的机率是 a，不发生的机率是 b，试问需要做多少次试验才能使该事件有可能发生，或者需要做多少次试验才能保证事件发生与否出现均等的机会.[6]

棣莫弗在其解答的开始部分分析到，如果需要 x 次试验，那么事件始终未发生的概率是 $\frac{b^x}{(a+b)^x}$.由于这种情况与事件在 x 次试验中是否至少发生一次的机会是均等的，所以该概率必然等于1/2，也就是说，x 必然满足方程

$$\frac{b^x}{(a+b)^x}=\frac{1}{2} \quad 或 \quad (a+b)^x=2b^x.$$

棣莫弗通过取对数的方法，很容易地给出了这一问题的解：

$$x=\frac{\log 2}{\log(a+b)-\log b}.$$

此外，他分析到，如果 $a:b=1:q$，即事件不发生与发生的机率之比为 q 比1，那么原始方程可以写作如下形式：

$$\left(1+\frac{1}{q}\right)^x=2 \quad 或 \quad x\log\left(1+\frac{1}{q}\right)=\log 2.$$

将 $\log\left(1+\frac{1}{q}\right)$ 展开为幂级数之后，棣莫弗得出结论，如果"q 是无限大，或与单位1比较而言相当大"，[7] 那么级数只需展开至首项 $1/q$，而问题的解则可以表示为 $x=q\log 2$ 或 $x=0.7q$.如此，为了解决德梅雷的问题，即需要将两个骰子掷多少次，才足以使两个六点同时出现与否的机会达到均等的程度，棣莫弗相当简单地取 $q=35$，因而 $x=24.5$.所以抛掷骰子的次数应该在24次和25次之间，这与惠更斯经过繁复的细致计算所得的结果完全相同.

人物小传	亚伯拉罕·棣莫弗(1667—1754)(Abraham De Moivre)
	棣莫弗生于法国巴黎以东约100英里的威特利(Vitry)的一个新教徒家庭.在11岁到14岁之间，他在色当的新教徒中级学校接受古典文学的教育，但当学校于1681关闭后，他先是转到索米尔(Saumur)再转到巴黎.在索米尔，他阅读了惠更斯的概率论著作；在巴黎他学习物理，并从欧几里得几何开始学习标准的数学课程.在南特法令(the Edict of Nantes)于1685年被废除之后不久，棣莫弗遭受了两年多的监禁生活.1688年4月，他重获自由，随即永久地离开法国，前往英国.正是在英国，棣莫弗掌握了牛顿的流数术并开始了自己的创造性工作.尽管在1697年被选入皇家学会，但他从未得到过大学职位.他的谋生手段除了做家庭教师，就是为赌徒或投机商解决机会游戏或年金保险中出现的问题.

棣莫弗经常借用无穷级数来计算概率.但比这种计算本身更为重要的是，他对二项式 $(a+b)^n$ 各项求和近似方法的详细论述.该方法问世于1733年，后以附录的形式被收入《机会学说》的第二、三版，并在首次出现时冠名为二项分布的正态近似(normal approximation to the binomial distribution).棣莫弗在其论述中体现了与伯努利相似的目标，即通过试验来估算概率："如假设某事件发生与否的可能性相同，尽管在3000试验中，该事件发生2000次而未发生1000次的情形缺少出现的可能，但这并不意味着它不会出现；而且一旦出现，随之产生的与平等相差悬殊的比率关系也必须被接受.由此，从试验中获取结论的思维应该更好."[8] 对于棣莫弗而言，与伯努利一样，计算

相关概率的方法有赖于对某些二项式系数的计算.他起初把自己严格限制在对等可能事件的研究中,以寻找在 n 次试验的半数情况下($n/2$ 次),此类事件发生的概率,即在 n 为较大的偶数时,二项展开式$(1+1)^n$ 的中项与所有项之和 2^n 的比率.他求得,当 n 增大时,该比率$\binom{n}{n/2}:2^n$ 趋近于 $\frac{2T(n-1)^n}{n^n\sqrt{n-1}}$,其中

$$\text{Log } T=\frac{1}{12}-\frac{1}{360}+\frac{1}{1260}-\frac{1}{1680}+\cdots=\frac{B_2}{1\cdot 2}+\frac{B_4}{3\cdot 4}+\frac{B_6}{5\cdot 6}+\frac{B_8}{7\cdot 8}+\cdots,$$

而 B_i 是伯努利数.

棣莫弗推导其计算结果的过程,显示了他对无穷级数和对数运算非常熟悉.他一开始便指出,中项 $M=\binom{n}{n/2}=n!/(n/2)!^2$,可以被表示为

$$M=\frac{(m+1)(m+2)\cdots(m+(m-1))(m+m)}{(m-1)(m-2)\cdots(m-(m-1))m},$$

其中 $n=2m$.接着,$\log M$ 可以被表示为各因子之商的对数,而这些对数中的每一个都可以展开为 $1/m$ 的幂级数.因此

$$\log M=\log\frac{m+1}{m-1}+\log\frac{m+2}{m-2}+\cdots+\log\frac{m+(m-1)}{m-(m-1)}+\log 2,$$

各对数如

$$\log\frac{m+1}{m-1}=\log\frac{1+\frac{1}{m}}{1-\frac{1}{m}}=2\left(\frac{1}{m}+\frac{1}{3m^3}+\frac{1}{5m^5}+\cdots\right),$$

而

$$\log\frac{m+2}{m-2}=\log\frac{1+\frac{2}{m}}{1-\frac{2}{m}}=2\left(\frac{2}{m}+\frac{8}{3m^3}+\frac{32}{5m^5}+\cdots\right).$$

然后,棣莫弗相当聪明地指出,这些幂级数的和可以通过垂直相加而非水平相加得以确定.因而,除了 $\log 2$ 那一项,幂级数的总和可以表示为如下各列和式的总和:

$$\text{col.}1=\frac{2}{m}(1+2+\cdots+s),$$

$$\text{col.}2=\frac{2}{3m^3}(1^3+2^3+\cdots+s^3),$$

$$\text{col.}3=\frac{2}{5m^5}(1^5+2^5+\cdots+s^5),$$

$$\cdots=\cdots$$

其中 $s=m-1$.以上各列和式由于涉及到整数列的幂和,故可以借助伯努利的公式,把它们转化为关于 s 的多项式来进行计算.棣莫弗将所有多项式的最高次项相加,得到一个幂级数,其和可以用有限项表示为$(2m-1)\log(2m-1)-2m\log m$.相似地,所有多项式的次高次项的和形成的幂级数相当于函数$(1/2)\log(2m-1)$.第三、第四……高次项的和很难计算,但棣莫弗证明了,当 m 趋向无穷时,这些和将变为$\frac{1}{12}$, $-\frac{1}{360}$,…考虑到特别项 $\log 2$,棣莫弗得出结论称,M 的对数为

$$\left(2m-\frac{1}{2}\right)\log(2m-1)-2m\log m+\log 2+\frac{1}{12}-\frac{1}{360}+\cdots,$$

而减去 $\log 2^n = \log 2^{2m} = 2m\log 2$ 之后,得到比率 $M:2^n$ 的对数为

$$n\log(n-1)-\frac{1}{2}\log(n-1)-n\log n+\log 2+\frac{1}{12}-\frac{1}{360}+\cdots.$$

因而,正如所言,$M:2^n = \dfrac{2T(n-1)^n}{n^n\sqrt{n-1}}$.

由于棣莫弗希望能够利用该比率进行具体计算,他借助于 $\log T$ 的级数形式证明了 $2T$ 近似于 2.168 = 221/125. 他还使用类似于此前所谈及的方法,在 m 很大的情况下,求得

$$\log m! = \sum_{k=1}^{m}\log k \approx \left(m+\frac{1}{2}\right)\log m - m + \log B \quad 或 \quad m! \approx Bm^{m+1/2}e^{-m},$$

其中 $\log B = 1 - \log T$,现今被称为斯特林(James Stirling,1692—1770) 公式. 斯特林曾求得 $B = \sqrt{2\pi}$,其计算方法有可能与棣莫弗的方法相类似,但他从沃利斯关于 π 的研究成果出发,随后便得出 $\log T = 1-\frac{1}{2}\log 2\pi$,或者 $T = e/\sqrt{2\pi}$. 棣莫弗知道,如果 n 很大,那么 $\dfrac{(n-1)^n}{n^n} = \left(1-\dfrac{1}{n}\right)^n$ 约等于 e^{-1},因此,他得出结论称 $(1+1)^n$ 的中项 M 与其总和 2^n 的比率等于 $2/\sqrt{2\pi n}$.

为了处理除中项以外的其他项,棣莫弗在某种程度上推广了自己的方法,并断言如果二项展开式 $(1+1)^n$ 中的一项 Q 从中项 M 起数为第 t 项,那么

$$\log\frac{M}{Q} = \left(m+t-\frac{1}{2}\right)\log(m+t-1)+\left(m-t+\frac{1}{2}\right)\log(m-t+1)-2m\log m+\log\frac{m+t}{m}$$

其中 $m = n/2$. 他再次基于以幂级数近似计算对数的方法,推断到,当 n 很大时,有

$$\log\frac{Q}{M}\approx -\frac{2t^2}{n} \quad 或 \quad Q \approx Me^{-(2t^2/n)}.$$

若用现代符号表示其含义,即是

$$P\left(X=\frac{n}{2}+t\right)\approx P\left(X=\frac{n}{2}\right)e^{-(2t^2/n)} = \frac{2}{\sqrt{2\pi n}}e^{-(2t^2/n)}.$$

棣莫弗考虑到 $Q = P\left(X=\frac{n}{2}+t\right)$ 的所有取值会形成一条曲线:"如果视二项展开式的各项为一系列竖直线段的长度,把这些线段摆放于一条直线上方且与之垂直,那么线段的上端点将描绘出一条曲线. 如此得到的曲线具有两个拐点,它们分别位于最大项对应点的两侧."[9] 该曲线现今被称之为正态曲线. 他算出,该曲线的两个拐点与最高点之间的距离为 $\frac{1}{2}\sqrt{n}$.

基于自己对二项展开式各项的近似计算,棣莫弗已经能够算出此类项的大数目的和,从而极为显著地改进了伯努利对于不确定性的量化方法. 为了计算

$$\sum_{t=0}^{k}P\left(X=\frac{n}{2}+t\right),$$

他将其近似为

$$\frac{2}{\sqrt{2\pi n}}\int_0^k e^{-(2t^2/n)}dt,$$

并通过幂级数展开再逐项积分的方法来估算该式中出现的积分值. 当 $k = \frac{1}{2}\sqrt{n}$ 时,展开的幂级数收敛速度相当快,这足以使他推知所求和等于 0.341 344,而且"如果可以无穷次地进行试验,那么对于发生与否具有等可能性的事件,其发生频率既不高于 $\frac{1}{2}n+\frac{1}{2}\sqrt{n}$ 也不低于 $\frac{1}{2}n-\frac{1}{2}\sqrt{n}$ 的概

率将是所给和数的两倍……,即 0.682 688."[10] 用现代的术语来说,棣莫弗已经证明了,当 n 足够大时,对称二项试验发生次数落入$\frac{1}{2}n$ 的$\frac{1}{2}\sqrt{n}$ 邻域的概率为 0.682 688. 棣莫弗随后计算了邻域半径为$\sqrt{n}$ 的各种分数倍时的相应概率值. 如此,"为了在各种特例中使用该方法,必需根据试验次数的平方根来估计事件发生与否的频率;而该平方根,…… 正像此前那样,将成为我们调控估计结果的模数(modulus)."[11] 对棣莫弗而言,$\sqrt{n}$ 是从中心起算的距离单位,因而估算概率的精度将会随着试验次数的增加而增加.

上述讨论仅适用于发生成败具有等可能性的事件. 但棣莫弗也确实勾画出一套方案,通过证明如何近似计算二项展开式$(a+b)^n$ 的各项,其中 $a \neq b$,以将自己的方法推广到更为普遍的情形. 利用这些普遍方法,人们可以计算出只需要很少次的试验,就可以满足伯努利在其例题中对精度的要求. 事实上,对伯努利要求进行 25 550 次试验的情形,在棣莫弗那里只需要 6498 次. 然而,棣莫弗自己只给出等概率情形的例题. 例如,他证明了 3600 次试验足以保障这种事件的发生频率介于 1770 和 1830 次之间的概率为 0.682 688,或者发生频率介于 1710 和 1890 次之间的概率为 0.998 74. 不幸的是,尽管棣莫弗的计算结果在实际上远比伯努利的精确,但他并没能很好地利用它们. 显然,他甚至没有意识到自己所发现的曲线,除了展示可以用作为衡量试验精度估计的尺度外,还具有其他重要作用. 不过,他的著作注定了要对这个世纪后来的发展产生深刻影响.

14.1.3 贝叶斯和统计推断

棣莫弗和伯努利的工作都没有获得直接应用,其原因之一在于,他们未能及时地解决应用所必需面临的一个问题,即统计推断:根据实验数据,某特定事件在次数确定的试验中发生了若干次,那么该事件发生的概率一般应该是多少?他们仅能说出观察频率与所给概率之间的近似程度. 第一位试图直接解决如何根据观察频率来推算概率问题的人是贝叶斯(Thomas Bayes,1702—1761). 他临终前完成的论文"对《机会学说》中一个问题的解答"(An Essay towards solving a Problem in the Doctrine of Chances) 在他去世三年后发表.

贝叶斯在其论文中一开始便叙述了这一基本问题:"给定某未知事件(指发生的概率未知) 发生以及未发生的次数,求在一次独立试验中该事件发生的概率介于两个指定概率值之间的可能性."[12] 借助于现代符号,如果 X 表示在 n 次试验中事件发生的次数,x 表示事件在一次独立试验中发生的概率,而 r 和 s 是两个给定的概率值,贝叶斯的目标是要计算出 $P(r<x<s \mid X)$,即已知 X 的情况下,x 介于 r 和 s 之间的概率. 接下来,贝叶斯从概率的定义出发,公理化地发展了他所必需的两条基本结论. 命题 3 称:"两个相继事件先后都发生的概率是一个比率,它由第一事件发生的概率,以及在第一事件发生的前提条件下,第二事件发生的概率复合而成." 命题 5 现今被普遍称为贝叶斯定理:"如果存在两个相继发生的事件,第二事件发生的概率是 b/N,两者都发生的的概率是 P/N,而且首先发现第二事件确已发生,据此,我猜想第一事件也已发生,其概率无疑等于 P/b."[13] 借用现代符号,用 E 表示第一事件,F 表示第二事件,命题 3 可以被表示为 $P(E \cap F) = P(E)P(F \mid E)$,即两事件都发生的概率是 E 发生的概率与在 E 发生后 F 也发生的条件概率的乘积. 而贝叶斯定理自身则可以被表示为 $P(E \mid F) = P(E \cap F)/P(F)$,即如果 F 已经发生,那么 E 发生的条件概率等于两事件都发生的概率除以 F 单独发生的概率所得商. 于是,贝叶斯的基本问题就是计算 $P(E \mid F)$,其中 E 指事件"$r<x<s$",而 F 指事件"某情形在 n 次试验中共发生 X 次". 为了能在其计算中运用命题 5,他因而需要一种方法来计算两个概率 $P(E \cap F)$ 和 $P(F)$.

贝叶斯当然知道伯努利的计算结果,即若事件发生的概率为 a,不发生的概率为 b,那么,在 n

$= p + q$ 次试验中,事件发生 p 次而不发生 q 次的概率为 $\binom{n}{q}a^p b^q$. 但是,伯努利仅能粗略地给出这些项之和的近似值,而棣莫弗主要考虑的是 $a = b$ 的等可能情形. 贝叶斯利用棣莫弗的思想,通过面积方法向这一问题发起了直接攻势. 他先是借用面积关系构建了概率模型:

> 设想正方形桌面……*ABCD*(图 14.1)可以如此构造并摆放水平,即若球 *O* 或球 *W* 被抛向它时,它们落在桌面上任何相等区域内的机会都是均等的.…… 假设球 *W* 先被抛出,通过其落点画一条直线 *ot* 平行于 *AD*,分别交 *CD* 和 *AB* 于 *t* 和 *o*;然后,球 *O* 被抛掷 $p + q$ 即 n 次,若它在一次单独抛掷中落入 *AD* 和 *ot* 之间,则称之为在一次独立试验中发生了事件 *M*.[14]

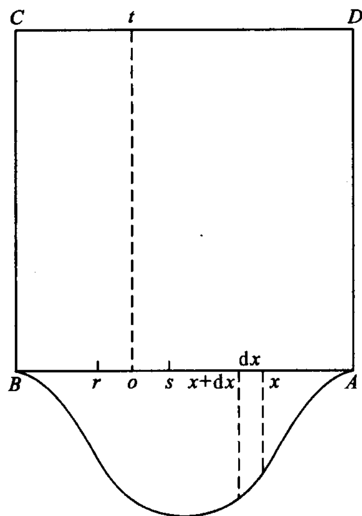


图 14.1 贝叶斯定理.

(为了以后叙述简单,规定 *AB* 的长度取值为 1.) 按照基本问题的说法,球 *W* 的位置决定着概率 x. 贝叶斯指出,点 o 落入点 r 和 s 之间的概率可以简记为 rs 的长度. 同样,在球 *W* 被抛出后,事件 *M* 的条件概率相当于 *Ao* 的长度.

为了计算 $P(E \cap F)$,贝叶斯利用了命题 3. 点 o 的任意给定的概率范围都可以用轴线 *AB* 上的一个区间来表示,如以端点 *A* 为起算点的区间 $[x, x + \mathrm{d}x]$. 由于一个特殊的 x 值表示了球落入 *ot* 右侧的概率,$1 - x$ 表示了球落入 *ot* 左侧的概率. 因而在 $p + q = n$ 次抛掷试验中,球将有 p 次落入右侧的概率可由 $y = \binom{n}{q}x^p(1-x)^q = \binom{n}{q}x^p(1-x)^{n-p}$ 给出. 贝叶斯在轴线 *AB* 的下方绘出了该方程所对应的曲线,并利用命题3推断出,球 *W* 落在区间 $[x, x + \mathrm{d}x]$ 的上方,而且另一球 p 次落入 *W* 右侧的概率,可以用介于 $[x, x + \mathrm{d}x]$ 下方和曲线上方之间的一块面积来表示. 因此,$P(E \cap F) = P((r < x < s) \cap (X = p))$ 可以表示为介于区间 $[r, s]$ 下方和曲线上方之间的一块面积,或者用现代符号表示为:

$$\int_r^s \binom{n}{p} x^p (1-x)^{n-p} \mathrm{d}x.$$

由于 $P(F) = P(X = p)$ 可以被看作 $P((0 < x < 1) \cap (X = p))$,所以,根据以上推断,$P(X = p)$ 可以表示为轴线 *AB* 下方和曲线上方之间的整块面积. 或者表示为

$$\int_0^1 \binom{n}{p} x^p (1-x)^{n-p} \mathrm{d}x.$$

于是,命题 5 暗示着

$$P(E|F) = P((r < x < s)|(X = p)) = \frac{\int_r^s \binom{n}{p} x^p (1-x)^{n-p} \mathrm{d}x}{\int_0^1 \binom{n}{p} x^p (1-x)^{n-p} \mathrm{d}x}.$$

据此,贝叶斯断定:"对我那里所谓的事件 *M*,从它在一定次的试验中发生与不发生的次数,而无需知道更多与它有关的其他信息,人们就足以猜测出它发生的概率处于什么范围之内. 而且,通过对那里提及的面积进行一般计算,人们还会看到这种猜测正确时的机率."[15]

虽然贝叶斯问题实际已从形式上获得解决,但在人们能够将其做为实用方法接受之前,仍有两大障碍需要克服.首先,贝叶斯的在桌面上滚球式的物理类比真的反映出了适用于这种理论的实际问题吗?未知概率 x 的实际取值真的与水平桌面上球的滚动相同吗?贝叶斯对该问题的解决,依靠了对这种理论之应用范围的有效限制.他只考虑如下情形,即对于任意给定的试验次数 n,所有可能的发生结果 $X = 0, X = 1, X = 2, \cdots$ 具有相同的概率,也就是说,对于那些事件"我没有理由认为在一定次数的试验中,它的发生会偏向于某个可能的次数,而非其他次数."[16] 然而,贝叶斯的这种特殊声明引发了对于现时以及他所提及情况的广泛争论.不了解给定情形下事件发生的概率值,难道就可以断定所有结果的出现具有等可能性吗?

其次,真的有人能够计算出贝叶斯公式中的积分吗?贝叶斯试图通过将这些积分展开为幂级数的方法做到这一点.分母中的积分实际上等于 $1/(n+1)$.而分子上的积分,尽管在 p 或 $n-p$ 都很小时不难给出近似,但在其他情况下却难以计算.贝叶斯的一位朋友普瑞斯(Richard Price, 1723—1791)向皇家学会递交了自己的论文,他找到了当 p 接近于 n 时一些特殊情况的解.例如,当 $p = n$ 时,相应的商为

$$\frac{\int_r^s x^n \mathrm{d}x}{\int_0^1 x^n \mathrm{d}x} = s^{n+1} - r^{n+1}.$$

因此,假设除了发生次数为 1 外,我们对于事件 M 别无所知.那么,事件 M 未知的发生概率 x 大于 1/2,即介于 1/2 与 1 之间的机率是 $1^2 - (1/2)^2 = 3/4$.类似地,如果 M 发生了两次,则 x 大于 1/2 的机率为 7/8;换句话说,事件发生概率大于或小于 1/2 的机率比为 7 比 1,从而事件发生的可能性远高于公平机率.在相同的情形下,未知概率 x 大于 2/3 的可能性仍然要高于公平机率.

14.1.4　拉普拉斯的计算

贝叶斯公式为解决统计推断的基本问题提供了契机.数年后,拉普拉斯(Pierre Simon de Laplace, 1749—1827)在此方面取得了更深层次的进展.1774 年,拉普拉斯利用与贝叶斯的相类似的原理,根据实验数据,在本质上得出涉及以积分计算概率的相同结果.拉普拉斯重新考虑了罐中抽签的问题.假设从一个白签所占比例 x 未知的罐中,抽得 p 个白签,q 个黑签.对于 x 的任何猜测值,拉普拉斯证明了如何计算 x 与 $p/(p+q)$ 的差小于任意值的概率,其中想要多小就有多小.他实际上证明的是

$$P\left(\left| x - \frac{p}{p+q} \right| \leqslant \varepsilon \mid X = p\right) \cong \frac{2(p+q)^{3/2}}{\sqrt{2\pi}\sqrt{pq}} \int_0^{\varepsilon} \mathrm{e}^{-[(p+q)^3/2pq]z^2} \mathrm{d}z \cong \frac{2}{\sqrt{2\pi}} \int_0^{\varepsilon/\sigma} \mathrm{e}^{-(u^2/2)} \mathrm{d}u,$$

其中 $\sigma^2 = pq/(p+q)^3$.为了证明无论取什么值,当 $p+q$ 逐渐增大时,该概率值趋近于 1,拉普拉斯必须求出积分 $\int_0^{\infty} \mathrm{e}^{-(u^2/2)} \mathrm{d}u$.利用欧拉的一个计算结果,他实际上已能够证明该积分等于 $\sqrt{\pi/2}$,从而建立起自己的结论.(注意,事实上,棣莫弗已经证明了这一点.)

人物小传	皮埃尔·西蒙·德·拉普拉斯(1749—1827)(Pierre Simon De Laplace)
	拉普拉斯(图 14.2)出生于诺曼底.他于 1766 年进入卡昂大学(the University of Caen)为将来成为一名牧师而做准备,但却在那里发现了自己的数学天赋,于是在 1768 年前往巴黎继续自己的学业.他遇到达朗贝尔并给其留下了深刻印象,在其担保下获得巴黎军事学院的一个数学职位,为这里满怀抱负的军校生们讲授初等数学.据传拿破仑于 1785 年在这儿参加并通过了他的考试.一系列数学论文很快如泉水般从他的笔下涌出,这使得他能够于 1773 年被选入科学院.法国大革命时期,他曾在度量衡委员会工作,但因作为共和党人缺少强硬的立场而最终被解职.他退休回到乡下,以求能够相对平静的从事研究工作. 拉普拉斯最深的造诣在天体力学领域.从 1799 年到 1825 年间,他著成自己的五卷本《天体力学》(Traité De Mécanique céleste),将微积分成功地应用于天体运动的计算,并证明了诸多结论,其中包括牛顿的引力定律为什么能够说明太阳系的长期稳定性.拉普拉斯在概率论方面也有重大贡献.他于 1812 年著成《概率的分析理论》(Théorie analytique des probabilités).尽管倍受拿破仑的尊敬,但他作为议会委员还是在 1814 年给拿破仑投了反对票,转而支持路易十八.拉普拉斯被授予侯爵的头衔并在死后被盛赞为"法国的牛顿".

当然,为了使计算更进一步,拉普拉斯还必须对任意 T 估算积分 $\int_0^T e^{-(u^2/2)}du$.他在 1785 年做到了这一点.其方法源起于该积分的两个不同的级数,其一对于较小的 T 收敛很快,而另一个则对于较大的 T 收敛很快.随后,他将自己的结论应用于一个真实的统计推断问题.在从 1745 年到 1770 年的 26 年间,巴黎出生了 251 527 位男孩和 241 945 位女孩.他以 x 表示男孩的出生率,用自己的分析方法进行了简单明了的计算,证实 $x \leqslant 1/2$ 的概率是 1.15×10^{-42},从而断言 $x > 1/2$ 具有"近乎必然性".他还利用从伦敦获得的类似数据来拓展自己的分析范围,以证明伦敦的男孩出生率近乎必然地高于巴黎的男孩出生率.



图 14.2 法国邮票上的拉普拉斯.

当统计推断中的实际问题获得解决之后,拉普拉斯又把自己的注意力转向了天文学.他对于概率论的更多贡献,部分得自对天文观测误差的分析,这将留待第 16 章讨论.

14.2 代数与数论

代数学与其他领域相比,在 18 世纪几乎没有取得什么大的新进展.仅有的几件重要成果不过是对早期工作的系统化概括,而且它们还是由那些在其他领域内更具影响力的数学家们完成的.这里将要介绍三部主要的代数学著作,一为牛顿所著(根据他从 1673 年到 1683 年间在剑桥的讲义改编而成,1707 年出版);一为麦克劳林所著(尽管有可能完稿于 18 世纪 30 年代,但直到 1748 年才出版);一为欧拉所著(1767 年出版).它们皆致力于将学生们引入代数领域并为其进一步的研究准备基础.牛顿的书流传至 19 世纪初期时,已经出现了拉丁文、英文、法文的许多版本.麦克劳林的书也再版过多次,最后一次是在 1796 年.欧拉的书在初版问世后的 50 年里,用六种语言重印了至少 30 次.这些著作反映了 18 世纪代数学的重要内容究竟是什么.

14.2.1　牛顿的《通用算术》

牛顿在剑桥讲授了十年的代数学，直到 1683 年，他才最终决定应该遵守卢卡斯教授职位的章程，从而在 1683—1684 年间的那个冬天补写了所有讲义，谨慎地标记下每份讲义应该被递交的日期，并按要求将其存入了学校的图书馆. 大约 20 年后，牛顿的继任者怀斯顿(William Whiston, 1667—1752) 对其进行了整理以备出版. 虽然牛顿对此并非完全满意，但它还是在 1707 年被公开出版.

牛顿的《通用算术》在开篇时所要求的知识水平非常初等，但在结束时给出的却是一个相当全面的课程，其中包含着代数方程解的许多引人关注的细节. 先来看看牛顿对加法与乘法的处理:

> 加法：在加数不是过分复杂的情况下，加法运算是不证自明的. 所以 7 和 9 的结果一看便明确，即 $7+9$ 等于 16，而 $11+15$ 等于 26. 但在比较复杂的情形下，加法运算的完成需要把各加数从大位到小位排列来分别计算各数位对应列的和.[17]
>
> 乘法：简单的代数项进行乘法运算，原则上乘以数则为数，乘以变量则为变量，而且在两个因子皆正或皆负时，乘积为正，否则为负.[18]

牛顿在此，或其他地方，都没有尝试去证明这条乘法法则的合理性，而仅限于陈述. 他同样没有证明自己提出的其他代数算法的合理性. 显然，明辩的理由对于他的听众或许还有读者而言并非必要. 所有必要的东西只是运算技术. 对此，牛顿给出了大量操作方法，既有关于数字的，也有关于代数表达式的. 他还讨论了解方程的基本方法，并花费大量时间来说明如何将各种问题代入代数学，其中包括许多来自几何学的问题. 他向人们展示了二次方程的求根公式，以及三次方程的卡尔达诺公式 —— 尽管对于后者，他认为"极为少用."

牛顿的许多"文字题" 很为人们所熟悉，今天的代数课本中仍有不少它们的翻版:

> 两个信使 A 和 B 相距 59 英里. 在一个早晨，他们相向而行以迎接对方. A 走完 7 英里需用 2 小时，B 走完 8 英里需用 3 小时. 现在 B 比 A 晚出发 1 小时，问：A 还须走多远才能与 B 相逢?[19]
>
> 若一个抄写员可以用 8 天时间誊写 15 张稿纸，那么在 9 天时间里誊写 405 张稿纸需要具有同等效率的抄写员多少名?[20]

在其著作的结尾部分，牛顿解决了许多更为复杂的问题，其中包括物理学和天文学上的一些问题. 他还发展了笛卡尔的符号法则、多项式方程系数与根的关系、多项式方程根的各整数次幂之和的计算公式. 但是，自 1707 年以后，牛顿不再将主要精力过多地投入数学研究，所以他从未使自己的这些工作变得精炼无瑕. 只有等他的后来者们麦克劳林和欧拉，吸收了他的思想并在其著作中修订了他的实例，才使其产生了更大影响.

14.2.2　麦克劳林的《代数学》

麦克劳林，像牛顿一样，认为代数学是"用特定的标记和符号进行运算的一种普遍方法. 这些标记和符号专门为此而设计且被证明相当方便. 它也被称为泛算术，基于与普通算术中类似的运算法则、相同的基本原理被构建而成."[21] 换言之，对于麦克劳林而言，代数学并不"抽象"，它只不过是被简单推广了的算术而已. 这样，为了理解代数就必需先搞懂算术. 所以，麦克劳林在其《代数学》开

篇,不仅给出了各种计算方法,还试图去解释它们背后的推理过程.例如,他在处理负数问题时写到,任何量都可以作为增加数或减少数之一被引入代数运算.作为这两种情形的例子,他引用了这样一些概念,如盈余和不足,一个人应得和应付的钱数,向右和向左所引直线,相对于地平线向上为升高和向下为降低.他认为,可以从相同性质的小量中减去大量,此时余数的性质会发生逆转.但人们只有在运算具有实际意义时才可以这样做,比如,从较少量的物质中减去更多量的物质就是不允许的.不过,麦克劳林一直认为负量并不比正量缺少现实意义,还因此而给出正量和负量参与运算的示例.特别地,为了展示这种量在相乘时符号法则成立的理由,他指出,由于 $+a-a=0$,且 $n(+a-a)=0$,但该乘积的首项 $+na$ 为正,故而其第二项必然为负.所以 $-a$ 乘以 $+n$ 为负.同理,由于 $-n(+a-a)=0$,该乘积的首项为负,则第二项 $(-n)(-a)$ 必然为正,且等于 $+na$.

麦克劳林在其著作的第一部分继续讨论类似主题,如分数运算、二项式的乘幂、多项式求根、级数求和等.他向读者们展示了 $(a+b)^n$ 展开式中各项的计算方法,其中 n 可以是正数或分数.关于 n 为分数时的计算自然会导致一个无穷级数.他说明了如何解线性方程和二次方程,并给出许多"应用问题"的实例.对于具有多个未知数的线性方程组,他证明了在方程个数与未知数个数同为 2 或 3 的情形下,可以借助替代的方法将其转换为只含一个未知数的线性方程来求解.他指出,如果未知数的个数多于方程个数,那么方程组将有无穷多组解.在相反的情形下,方程组可能没有任何解.但他针对这两种情形没有给出任何例题.

然而,麦克劳林的确为方程组的消元方法给出一条自称的"普遍定理",也就是现今所谓的克拉默法则.它得名于瑞士数学家克拉默(Gabriel Cramer,1704—1752),他曾于 1750 年在一本论述曲线的书中使用过它.如果

$$ax+by=c,$$
$$dx+ey=f,$$

那么,从第一个方程中解出 x 并代入第二个方程,可求得

$$y=\frac{af-dc}{ae-db},$$

以及 x 的一个类似答案.对于三次方程组

$$ax+by+cz=m,$$
$$dx+ey+fz=n,$$
$$gx+hy+kz=p,$$

先从每个方程中求出 x,从而将其简化为只有两个未知数的方程组,再利用前边的法则求得

$$z=\frac{aep-ahn+dhm-dbp+gbn-gem}{aek-ahf+dhc-dbk+gbf-gec}.$$

在给出该答案之余,麦克劳林明确了那条普遍法则,即解的分子由 x 和 y 的系数以及常数项的各种乘积组成,每个乘积只含有各方程中的一个系数,而解的分母则由三个未知数的系数的乘积组成.他还解释了如何确定每一项乘积的符号.此外,他解出方程中的 y 和 x,并指明确定其值的普遍法则与确定 z 值的法则相类似.特别地,方程组解的三个表达式都具有相同的分母.麦克劳林甚至还将这一法则延伸到了含有四个未知数的方程组,可惜未能提及任何更进一步的推广.

当然,麦克劳林给出的解的分子和分母实际就是现今为人所知的"行列式".但利用系数的此类组合作为工具以求解线性方程组的有关思想却在较早以前就有出现.1693 年,莱布尼兹在给洛必达的一封信中就曾提到过类似想法,他甚至还发明了一种以数字为标识来指代方程组系数的方法.在世界的另一边,比此早十年,日本数学家关孝和就在一部手稿中描述了行列式的应用,并详细展示

了如何借助于图表来确定给定项的正负号.

人物小传	关孝和(1642—1708)
	关孝和,因其名字的日文字母具有不同读法,或称 Seki Takakazu,或称 Seki Kōwa.在距东京西北 50 英里的藤冈小镇,关孝和出生于一位封官候的扈从之家.在位于现山梨县境内的甲府,他为两位幕府将军作会计师.后来,迁居东京从事相同工作.尽管关孝和很少发表文章,但有迹象表明,在他的众多手稿中,他理解方程理论的许多基础知识,其中包括多项式方程具有与其次数同个数的解的概念.而其工作被关孝和研究过的中国数学家们,却只满意寻求一个解.1683 年,关孝和在一部讨论建立并求解方程的著作中引进了行列式.他所用的图表显示了决定乘积因子及其符号的方法,其中之一在插图中部分可见(图 14.3).



图 14.3　日本邮票上的关孝和.

麦克劳林著作的第二部分以严密的组织结构讨论了截止他那个时代已经被发现的多项式方程的求解方法,这不仅包括求解三次方程的卡尔达诺法则和求解二次方程的费拉里法则,而且包括笛卡尔的符号法则和牛顿的求方程数值近似解的方法.他指出,方程的形成过程,将类似于 $x-a=0$ 或其他次数低于给定次数的方程相乘而得到,说明了没有哪个方程的根的个数会超过最高次幂的次数.而且,"根将成对地变得不可能(复数根)" 所以"奇数次方程总有一个实数根."[22] 他然后讨论了求首项系数为 1 的多项式方程的整数根的一般程序:检查常数项的所有因子是否为可能的根,如果有一个这样的根 α 被找到,将多项式除以 $x-\alpha$ 就可以降低方程的次数.

麦克劳林著作的结论部分讨论了代数方法在几何学上的应用,以及相反地,几何手段在解方程方面的应用.他指出,代数与几何在应用方面最主要的差别在于,对于前者,即使虚数根也可以被明确地表示出来,而对于后者,这样的量从来不会出现.麦克劳林在这一部分中,还讨论了运用圆来构造二次方程解,以及运用圆锥曲线来构造三次、四次方程解的详细法则.虽然文中几乎没有数学上的新东西,但麦克劳林的著作对于 18 世纪的学生们而言仍不失为当时代数学的一个很好的导引.

14.2.3　欧拉的《代数学引论》

对代数学更好的介绍或许应该算是欧拉的《代数学引论》(Vollständige Anleitung zur Algebra).欧拉像麦克劳林一样,在其著作的开始部分给出了该主题的一个定义:"所有数学科学的基础必然潜藏于对数的科学的全面论述,以及对各种可能的计算方法的精确考察当中.数学的这一基本部分被称为分析,或者代数.因此,在代数学中,我们只研究代表着各种量的数,而无需考虑这些量的实际背景."[23] 稍后,他又在文中将这一界定变得更加详细而精确:代数学是"教授如何依靠已知量来计算未知量的科学."[24] 他指出,在该定义下,即使通常意义上的两个已知量的加法运算也可以被认为是适宜的,所以第二个定义或许更为普遍,且在实际上包含了第一个定义.

欧拉从正负数的运算开始自己的讨论.与麦克劳林相比,他关于乘法的讨论多少有点不够规范:"让我们从 $-a$ 乘以 $+3$ 开始.现在,由于 $-a$ 可能被认作负债值,显然,如果我们三次承担这么多债务,负债值将变成原来的三倍.所以,要求的乘积为 $-3a$."[25] 欧拉随即对此作出明显推广,即 $-a$ 乘以 b 等于 $-ba$ 或 $-ab$.接着是两个负数相乘的情形.他在此只是简单地说,$-a$ 乘以 $-b$ 不可能等同于 $-a$ 乘以 b,即 $-ab$,所以其结果必然等于 $+ab$.

在讨论完其他各种运算之后,欧拉引入了虚数的概念:

既然可以被实际构想出来的所有数要么大于0,要么小于0,要么自身等于0,显然,一个负数的平方根将无法被排入这些可能的数当中,因而,我们只有称它为一个不可能的量.如此,我们可以被引向一类特别的数,它们在本质上是不可能的量,由于它们只存在于想象之中,因此常常被称为虚构的量(imaginary quantity).所有此类表达式,如$\sqrt{-1}$,$\sqrt{-2}$,⋯ 都是不可能的数,或称虚数,这是因为它们都是负数的平方根;⋯⋯ 尽管如此,这些数仍出现在我们的头脑中;它们存在于我们想象里,我们对它们有着充分的理解;我们还知道$\sqrt{-4}$是这样一个数,它乘以自身其结果为 -4;正是由于这个原因,没有什么可以阻止我们使用这些虚数,并将其用于计算之中.[26]

奇怪的是,欧拉并没有意识到在此类运算中可能会有问题出现.尽管已经注意到$\sqrt{-4}\times\sqrt{-4}=-4$,但他在稍后又写到,平方根乘积的普遍法则意味着$\sqrt{-1}\times\sqrt{-4}=\sqrt{(-1)(-4)}=\sqrt{4}=2$.

欧拉接下来在文中讨论了对数、无穷级数,以及二项式定理.他定义对数的方法如同他在《无穷小分析导论》中所做的那样:如果 $a^b=c$,那么 b 是c 的以a 为底的对数.对数随后在论述有关复利计算的章节得到应用.无穷级数是按照除法的思想被引入的,第一个例子就是$1/(1-a)=1+a+a^2+a^3+\cdots$.虽然欧拉没有利用相同篇幅来讨论收敛性,但他仍断言:"有足够的理由相信该无穷级数的值与该分式的值相同."[27] 随后,他给出一些例子以说明自己的陈述将"很容易被理解".如此一来,如果 $a=1$,分式的值等于1/0,"是一个无穷大的数",而级数变为$1+1+1+\cdots$,也是无穷大,从而该结论得到佐证.欧拉推断,如果 a 的取值比1小,那么"一切都变得更加易于理解".在这种情形下,"级数的项被展开得越多,[分式与级数间的]差就越小;因而,如果级数被无穷续写下去,其和与分式值之间将无所差."[28]

欧拉著作的最后一部分致力于讨论在牛顿和麦克劳林的著作中不曾出现的主题,即不定方程的解.该部分解决的许多问题实际上是丢番图的《算术》中的问题.但欧拉如同一个世纪以前的费马一样,通常给出的是问题的一般解,而非带有古希腊代数学家特征的单独解.作为例子,考虑如下问题,这实际上就是丢番图的问题 II-11:

问题2 求数 x,以使其分别加任何两个数,如4和7,在这两种情形下其结果都是一个平方数.[29]

丢番图曾用双重方程的方法(method of the double equation)求解此题.欧拉则使用了不同的技巧.设 $x+4=p^2$,得到 $x+7=p^2+3$ 是一个平方数,其根为 $p+q$.再设 $q=r/s$,那么 $p^2+3=p^2+2pq+q^2$,或者,$p=(3-q^2)/2q$,或者最终得到

$$x=p^2-4=\frac{9-22q^2+q^4}{4q^2}=\frac{9s^4-22r^2s^2+r^4}{4r^2s^2}.$$

欧拉然后指出,选择 r 和s 而得到的任何整数都确定x 的一个解.

然而,在论述不定方程的这一章节中,大部分内容的重点在于讨论一般方法,而非特殊问题.对于形如$p(x)=y^2$的方程,其中$p(x)$是2、3或4次多项式,欧拉特别论述了求其有理解或整数解的技巧.作为特例,他考虑第6章曾讨论过的方程 $Dx^2+1=y^2$的整数解.费马曾宣称得到了该解,但却被欧拉错误的归功于英国数学家佩尔(John Pell,1610—1685).欧拉在此没有给出一般求解方法,而是针对每种情形给出一种具体方法.随后,他以一个数表结束了自己的讨论.在该数表中,罗列着D从2到100取值时方程的解.虽然欧拉没有证明对于任何 D,方程的解都存在,但在1766年,拉格朗日给出了这样的证明,并被作为附录收入《代数学引论》稍晚的一个版本.

14.2.4 欧拉与线性方程组

在《代数学引论》中,欧拉没有特别具体地讨论线性方程组的求解,而只是简单地说明,在解方

程组时,可以先把一个未知数用其他未知数表出,然后进行替换,如此可将方程组降为含有更少方程和更少未知数的方程组.

另一方面,截止 1750 年,欧拉一直在寻找求解方程组的更为一般的思想方法.在该年的一篇论文中,他专注于解决克拉默悖论(Cramer's paradox).该悖论基于 18 世纪人人都确信的两个命题:(1) 一条 n 次代数曲线由其上的$n(n+3)/2$个点惟一确定.(2) 两条 n 次和 m 次的代数曲线有 nm 个交点.第一个命题的结论来自初等组合学.一条 n 次代数曲线,即关于两个变量的 n 次多项式所描述的曲线,基本上具有1个零次项系数,2个一次项系数,3个二次项系数,4个三次项系数,如此等等.但由于可以用任何一个系数去除其他系数,所以,"独立" 的系数的总个数为

$$\sum_{i=1}^{n+1} i - 1 = \frac{(n+1)(n+2)}{2} - 1 = \frac{n(n+3)}{2},$$

如此多个数的点恰好是决定一条代数曲线所必需的.事实上,正是根据曲线上的某些已知点来确定曲线系数的问题引导克拉默发现了自己的克拉默法则,并简明地揭示了克拉默悖论.对于第二个命题的结论,虽然已知代数曲线的这些交点可能会有重复或者是虚假的,但仍可以找到很多例子,其中所有 mn 个交点是实实在在且彼此不同的.克拉默悖论来自对 $n \geqslant 3$ 时的情形的考虑.根据第二个命题,两条 n 次代数曲线存在n^2 个公共点,而第一个命题却暗示着只需 $n(n+3)/2$个点(少于 n^2 个)就能惟一确定一条曲线.

欧拉讨论了克拉默悖论并且发现,第一个命题的结论建立于 n 个 n 元线性方程应该确定着惟一一组n维解的事实,故在不加限制条件的情形下未必成立.当时的人们非常坚信 n 个方程确定了 n 个未知数,以致于早些时候没有人真正耐心地讨论过它不会发生的情形.如前所述,麦克劳林简要地讨论了方程个数与未知数个数不等的情形,但并未提及目前这种情形出现的可能性.

在他的论文中,欧拉讨论了各种例子,但未能提出一个明确的定理.例如,他指出 $3x - 2y = 5$ 和$4y = 6x - 10$不能确定两个未知量的值,这是因为解出 x 后进行替换,则关于 y 的方程将变成恒等式,据以无法求出任何值.他还给出一个含有4个4元方程的方程组,以其他变量表示出其中的两个变量,代入另两个方程后,再次出现恒等式,使得后两个变量的值无法确定,从而这 4 个方程也不能确定 4 个变量的值.有鉴于此,他得出结论称,如果要说 n 个方程足以确定n 个未知量的值,那么就必须加上约束条件,即这些方程如此相异以致于没有一个已被其他方程所"包含".虽然欧拉没有明确地界定"包含" 的含义,但可以看出,他至少在直觉上理解了方程组"秩" 的概念.

为了解决克拉默悖论,欧拉最后指出:"当两条四次曲线相交于16个点时,由于仅需 14个点,在它们引出的方程互异时,就足以确定具有该次数的一条曲线,所以在由这 16 个点引出的方程中总有三个或更多方程已被包含于其他方程之中.这样一来,这16个点并不能比13个,或12个,或更少的点所确定的东西更多,而且为了完全确定整条曲线,必须再在这 16 个点中加入一个或二个其他点.[30]

虽然欧拉解决了这个直接悖论,但数学家们仍花费了一个多世纪的时间来完全理解不定方程组或不相容方程组的概念.我们将在第 15 章继续对这些概念的讨论.

14.2.5 欧拉与数论

欧拉一生中还研究并解决了许多有趣的数论问题,其中有些问题是由费马提出或从费马已经解决的问题中派生出来的.如在 1749 年,欧拉证明了费马的一个猜想,即每个形如$4n+1$ 的素数可以被表示为两个平方数之和.1773 年,欧拉在数年研究的基础上,证明了每个整数可以表示为不超过四个的平方数的和.(拉格朗日已于三年前证明了该结论;欧拉的证明是对自己早期有关两个平方数之和证明的推广.)然而,我们在此仅讨论欧拉对于同余的详细研究、对于费马小定理的推广,

以及对于二次互反律的发现.

大约在1750年前后,欧拉开始着手撰写一部关于数论的初等论著,但在完成了16章之后,便将其搁置一边.他死后,这部手稿才被发掘出来并最终于1849年获得出版,名为《论数的学说》.该论著的前几章主要是计算数论函数,如表示整数 n 因子个数的函数 $\sigma(n)$,以及表示小于 n 且与之互素的整数个数的函数 $\varphi(n)$.然而,整个论著最为重要的部分却开始于第5章,欧拉在此论述了关于数 d 的同余概念.数 d 就是如今所谓的模.欧拉定义 a 关于 d 的剩余为除法运算中 a 除以 d 的余数 r,即 $a = md + r$.他指出,有 d 种可能的余数,故而所有整数被划分为 d 个类,每个类由具有相同指定余数的数组成.例如,除数为4的除法运算将整数划分为4类,每类中的数分别具有 $4m$, $4m+1$, $4m+2$ 和 $4m+3$ 的形式.给定类中的所有数都被其视为"等价".欧拉进一步说明,可以在这些类上定义运算.如此,若 A 和 B 分别属于 α 和 β 的剩余类,那么 $A+B$, $A-B$, nA 以及 AB 分别位于 $\alpha+\beta$, $\alpha-\beta$, $n\alpha$ 和 $\alpha\beta$ 所在的剩余类中.按照现代术语,欧拉论证的是,将整数归入其"剩余类"的函数是一个环同态.事实上,环理论的最终诞生正是基于这些思想方法.

同样,欧拉关于算术级数 $0, b, 2b, \cdots$ 中各剩余的讨论也显然具有群论的基本思想.欧拉证明了如果模 d 和数 b 相对互素,那么该级数包含着来自所有 d 个不同剩余类中的元素.从而 b 有一个关于 d 的"逆元素" p,使得 pb 的剩余为1.另一方面,如果 d 和 b 的最大公因子 $g > 1$,那么在该级数中只出现 d/g 个不同剩余类中的元素,且类似的逆元素将不存在.例如,2的倍数集包含了关于模9的9个不同剩余类中的元素,而且5就是2的逆元素.而3的倍数集仅包含关于模9的3个不同剩余类中的元素,而且不存在3的逆元素.

欧拉继续着这一研究路线,他考虑几何级数 $1, b, b^2, b^3, \cdots$,其中 b 与 d 互素.该级数中不同类剩余的个数 n 不会比 $\mu = \varphi(d)$ 还多.欧拉指出,数 n 是大于1且使 b^n 的剩余为1的最小数.这是因为,一旦乘幂次数达到 n,级数中的后续项将与前面的某项具有相同的余数.为了说明 n 是 μ 的因子,他用到一个后来在群论中相当标准的证明,事实上,即是考察在关于 d 的剩余的乘法群中,b 的乘幂的子群的同余类,并证明群的阶可以被其子群的阶整除.欧拉首先说明,若 r 和 s 分别是 b^{ρ} 和 b^{σ} 的剩余,那么,rs 是 $b^{\rho+\sigma}$ 的剩余,同样,r/s 也是一个剩余.因此,如果 r 是一个剩余,且 $x < d$ 是一个非剩余(与 d 互素,但非幂级数中各项的剩余),那么,xr 也必定是一个非剩余.所以,如果 $1, \alpha, \beta, \cdots$ 形成全体 n 个剩余的集合,那么 $x, x\alpha, x\beta, \cdots$ 就形成含 n 个不同非剩余的集合.由于不包括在后一集合中的其他任何非剩余也可以引出一个集合,它含有异于第一个集合中元素的 n 个非剩余,欧拉得出结论称,对于一些整数 m 有 $\mu = mn$ 成立,随之有 $b^{\mu} = b^{mn}$ 除以 d 后的余数为1,或者说,$b^{\mu} - 1$ 能被 d 整除.当 d 是一个素数 p 时,该结论有一个特例,那就是费马小定理.

在欧拉生命中的很长一段时间里,乘幂的剩余问题一直吸引着他的注意,而且他在从事计算时经常要用到它们.所以,手稿第7章末尾提供了一张表格,列出了数的乘幂及其关于模 d 从2到13时的所有剩余.欧拉还广泛地计算了形如 $x^2 + ny^2$ 的表达式的素因子,并试图寻求哪些素数可以被表示为如此形式.在1751年,他发表了一篇论文,其中仅包含了 n 取16个正值和18个负值时相应的此类结果.到1783年时,这些计算终于引导欧拉做出了等价于二次互反律的理论陈述.

欧拉称 $p \neq 0$ 是一个关于素数 q 的**二次剩余**是说,如果存在 a 和 n,使得 $p = a^2 + nq$,即若 $x^2 \equiv p \pmod q$ 有解.注意,作为关于 q 的二次剩余的条件仅依赖于关于 q 的剩余类 p.例如,1,4,9, $5 \equiv 4^2$ 和 $3 \equiv 5^2$ 都是关于11的二次剩余,而2,6,7,8以及10都是非剩余.在其1783年的论文中,欧拉首先证明了当 $q = 2m+1$ 是一个奇素数时,正好存在 m 个二次剩余,从而也就有 m 个非剩余.更进一步,他证明了两个二次剩余的积与商也是二次剩余.他接着证明了当 q 是形如 $4n+1$ 的数时,-1 是关于 q 的剩余,而当 q 是形如 $4n+3$ 的数时,-1 是非剩余.在文章的结尾处,他写下四个猜想.它们

是关于对两个相异的奇素数 q 和 s，一个对于另一个是或不是二次剩余的条件.具体地可以表述如下：

1. 如果 $q \equiv 1 (\bmod 4)$ 且 q 是关于 s 的二次剩余，那么，s 和 $-s$ 都是关于 q 的二次剩余.

2. 如果 $q \equiv 3 (\bmod 4)$ 且 $-q$ 是关于 s 的二次剩余，那么，s 是关于 q 的二次剩余而 $-s$ 不是.

3. 如果 $q \equiv 1 (\bmod 4)$ 且 q 不是关于 s 的二次剩余，那么，s 和 $-s$ 都是关于 q 的非剩余.

4. 如果 $q \equiv 3 (\bmod 4)$ 且 $-q$ 不是关于 s 的二次剩余，那么，$-s$ 是关于 q 的二次剩余而 s 是非剩余.

欧拉没能在 1783 年证明这些结论. 勒让德在其 1785 年的论文和 1798 年的课本《数论研究》(Essai sur la théorie des nombres) 中，以一种稍微不同的方式重新叙述了这些结论，然而在这两处都只有一个不很完善的证明. 第一个完整的证明由高斯在其 1801 年的名著《算术研究》中给出. 我们将其留待第 15 章讨论.

14.2.6 拉格朗日与多项式方程求解

在 18 世纪，代数学中受到关注的还有多项式方程求解. 事实上，许多数学家试图通过推广卡尔达诺和费拉里的方法来代数地求解五次或更高次多项式方程，但均以失败而告终. 直到 1770 年，拉格朗日发表《方程代数解之反思》(Réflexions sur la théorie algébrique des équations) 一文，才终于开创了此项工作的一个新局面. 他详细回顾了多项式方程的早期解法，以寻找求解三次和四次方程的方法之所以奏效的原因. 他虽然没能找到适用于更高次方程的类似解法，却针对此类方程概括出了一套他自认为终将奏效的新原理.

拉格朗日首先从卡尔达诺的程序入手，开始系统研究三次方程 $x^3 + nx + p = 0$ 的解法. 设 $x = y - (n/3y)$，将该三次方程转化为六次方程 $y^6 + py^3 - (n^3/27) = 0$，再设 $r = y^3$，进一步将其转化为二次方程 $r^2 + pr - (n^3/27) = 0$. 该二次方程有两个根，r_1 和 $r_2 = -\left(\frac{n}{3}\right)^2 \frac{1}{r_1}$. 然而，一反卡尔达诺以 r_1 与 r_2 的实立方根之和为原方程根的做法，拉格朗日知道，方程 $y^3 = r_1$ 与 $y^3 = r_2$ 都有三个根. 于是 y 有六个可能的值，即 $\sqrt[3]{r_1}$，$\omega\sqrt[3]{r_1}$、$\omega^2\sqrt[3]{r_1}$、$\sqrt[3]{r_2}$、$\omega\sqrt[3]{r_2}$ 与 $\omega^2\sqrt[3]{r_2}$，其中 $\omega = (-1+\sqrt{-3})/2$ 是方程 $x^3 - 1 = 0$ 或者说 $x^2 + x + 1 = 0$ 的一个复数根. 所以，拉格朗日能够证明，原方程的三个不同的根可以如下给出：

$$
\begin{aligned}
x_1 &= \sqrt[3]{r_1} + \sqrt[3]{r_2} \\
x_2 &= \omega\sqrt[3]{r_1} + \omega^2\sqrt[3]{r_2} \\
x_3 &= \omega^2\sqrt[3]{r_1} + \omega\sqrt[3]{r_2}.
\end{aligned}
$$

随后，拉格朗日指出，与其将 x 视为 y 的函数，还不如将问题反过来考虑，这是因为只要能从他所谓的约化方程中解出 y 值，即可促使原方程获解. 按照他的想法，接下来要做的即是用原方程的解去表示那些 y 值. 如此以来，拉格朗日指出，y 的任何六个值都能够以 $y = \frac{1}{3}(x' + \omega x'' + \omega^2 x''')$ 的形式表出，其中 (x', x'', x''') 是 (x_1, x_2, x_3) 的某些置换. 正是这种对方程根之置换概念的引入，无论为拉格朗日自己的方法，还是为下个世纪其他人将要用到的方法，都奠定了坚实的基础.

在三次方程的情况下，存在着许多值得注意的思想方法. 首先，x_i 的六个置换可以导出 y 的六个可能值，从而显示出 y 满足一个六次方程. 其次 y 的表达式的置换可以被分为两组. 其中有一组包含了恒等置换和交替取所有三个 x_i 值的两个置换；而另一组包含只在两个 x_i 间交替取值的三个

置换.(用现代术语讲,具有三个元素的集合上的置换群可以被分为两个陪集.)例如,如果 $y_1 = \frac{1}{3}(x_1 + \omega x_2 + \omega^2 x_3)$,那么,在第一组中的两个非恒等置换将把 y_1 分别变为

$$y_2 = \frac{1}{3}(x_2 + \omega x_3 + \omega^2 x_1) \quad 和 \quad y_3 = \frac{1}{3}(x_3 + \omega x_1 + \omega^2 x_2).$$

从而,

$$\omega y_2 = \omega^2 y_3 = y_1 \quad 且 \quad {y_1}^3 = {y_2}^3 = {y_3}^3.$$

类似地,如果第二组中的置换为 y_4, y_5 和 y_6,则有 ${y_4}^3 = {y_5}^3 = {y_6}^3$.从而由于 $y^3 = \frac{1}{27}(x' + \omega x'' + \omega^2 x''')^3$ 仅有两个可能的值,故关于 y^3 的方程是二次的.最后,y 所满足的这个六次方程具有原方程系数中的有理数系数.拉格朗日还考虑了其他一些求解三次方程的方法,并发现在各类情形下都同样潜在着这种思想.每种情形下的三个根都能导出一个有理表达式,它在六个可能的置换下仅有两个值,这说明该表达式满足一个二次方程.

拉格朗日随后研究了四次方程的解.在求解方程 $x^4 + nx^2 + px + q = 0$ 时,费拉里的方法是给方程两边同时加上 $2yx^2 + y^2$,然后重新整理,并确定一个 y 值使得新方程

$$x^4 + 2yx^2 + y^2 = (2y - n)x^2 - px + y^2 - q$$

的右端是完全平方数.在对方程两端开方之后,他能够求解导出的二次方程.上述方程右端为完全平方数的条件是

$$(2y - n)(y^2 - q) = \left(\frac{p}{2}\right)^2 \quad 或 \quad y^3 - \frac{n}{2}y^2 - qy + \frac{4nq - p^2}{8} = 0.$$

所以,约化方程是一个三次方程,从而自然是可解的.解出 y 的三个值之后,拉格朗日像在前一种情形下那样,证明了它们中的每一个都是原方程四个根 x_1, x_2, x_3, x_4 的有理函数的一个置换.事实结果是 $y_1 = \frac{1}{2}(x_1x_2 + x_3x_4)$,而 x_i 的 24 种可能的置换仅能导出该表达式的三种不同值,即 y_1,$y_2 = \frac{1}{2}(x_1x_3 + x_2x_4)$ 和 $y_3 = \frac{1}{2}(x_1x_4 + x_2x_3)$.因此,该表达式必然满足一个三次方程,其系数是原方程中的有理系数.

在对三次和四次方程求解方法的研究基础上,拉格朗日做好了推广的准备.首先,从对三次方程的讨论中可以清楚地看到,对于形如 $x^n - 1 = 0$ 的方程的根的研究很重要.当 n 为奇数时,拉格朗日可以证明所有的根能够表示为其中一个的乘幂.特别地,如果 n 是素数,$\alpha \neq 1$ 是其中一个根,那么对于所有的 $m < n$,α^m 就是所有根的生成元.其次,拉格朗日意识到为了解决 n 次方程的问题,他需要找到一种方法来确定约化方程,其次数 $k < n$.原方程根的某些函数必须满足这样的约化方程,而且这些函数在考虑到根的 $n!$ 种置换时仅有 k 个值.由于根的相对简单的函数并不可用,拉格朗日便试图寻找一个更为普遍的法则来推求这样的函数以及它们应当满足的方程的次数.

拉格朗日指出,如果约化方程的根为 $f_1, f_2, \cdots, f_k$,其中 f_i 是原方程 n 个根的一个函数,那么,约化方程可以通过 $(t - f_1)(t - f_2)\cdots(t - f_k) = 0$ 给出.尽管他并不能证明该方程的次数通常小于 $n!$,但他能够证明其次数 k,也就是在变量的置换意义下 f 所取不同值的个数,总能整除 $n!$.根据这一陈述,人们可以体会到拉格朗日的定理实际上给出的是群的阶可以被其任何子群的阶整除.然而,拉格朗日从未将置换作为运算"群"进行处理.不过,拉格朗日继续证明了根的函数是如何相互关联着.他证明,如果在根的所有置换中,那些能使某个函数 u 不发生变化的置换却可以使另一个函数 v 取得 r 个不同的值,那么 v 将是某 r 次方程的一个根,该方程的系数是 u 和原方程中系数的有理

函数.例如,在三次方程 $x^3+nx+p=0$ 中,表达式 $v=\frac{1}{27}(x_1+\omega x_2+\omega^2 x_3)^3$ 在根的六种置换下取两个值,而 $u=x_1+x_2+x_3$ 在这些置换下无变化.那么 $v^2+pv-(n^3/27)=0$ 就是 v 所满足的方程.(注意此时 $u=0$.)

拉格朗日大概很想通过使用这一法则来求解一般的 n 次多项式方程.也就是说,他可能会从根的一个对称函数开始,如 $u=x_1+x_2+\cdots+x_n$,它在 $n!$ 个置换下无变化.然后找到一个函数 v,它在这些置换下取 r 个不同的值.如此一来,v 将成为一个 r 次方程的根,该方程的系数是原方程中的有理系数(因为给定的对称函数 u 是这些系数之一).如果该方程能够被求解,那么,他就能找到一个新函数 w,假如它在那些使 v 不发生变化的置换下取 s 个不同的值,则 w 就应该满足一个 s 次的方程.他会一直这样做下去,直到找到函数 x_1.可惜的是,拉格朗日并没有找到一种普遍方法,来确定这些中间函数并使它们具有根据已知方法可求解的形式.他被迫就此放弃了这种探索,然而,他的工作却为 19 世纪关于方程代数解的所有研究打下了扎实的基础.相关内容将留待第 15 章继续讨论.

14.3 几何学

18 世纪的几何学既与代数学保持着依靠解析几何方法所建立起来的联系,又通过无穷小方法在曲线和曲面研究中的应用连接着分析学.此外,对于欧几里得平行公设的持续讨论仍吸引着人们浓厚的兴趣.在考虑几何学的以上各方面之前,我们还是先来谈谈一本介绍该领域的初级教材.

14.3.1 克莱罗与《几何原理》

克莱罗于 1741 年发表的《几何原理》(Éléments de géométrie)是 18 世纪重要的几何课本之一,它显示出其作者深信该学科的初学者们应当以一种他所谓的"自然" 方法来学习其基本原理.克莱罗写到:"我有意回到那些可能引起几何学诞生的问题上,而且试图通过极其自然的方法来发展几何学的基本原理,以使人们可以假想它正如其发明者最初所得到的一样,需要尽量避免的是这些发明者们可能不得不走过的弯路."[31] 他的课本对于几何教学相当有影响,一直流传到 19 世纪,历经了 11 个法文版本,并被译成瑞典语、德语和英语.

克莱罗认为,几何学应该从区域的度量开始.毕竟,几何学的名称(geometry) 本身就与土地丈量有关.因此,他呈现给学生们的首先是这方面的基本思想,而非欧几里得的公理和定义.他计划基于对测量基本法则的类推来发展更为复杂的思想方法,并贯穿始终地展示人们天生的好奇心如何驱使他们能够解决新问题和发现新概念,如此以期望在自己的读者中鼓励这种发现精神.他意识到自己可能会因不够"严格" 的证明而遭受批评,但仍坚持认为没有必要用抽象的推理去证明那些被判断力良好的人所认可的结论.

克莱罗的自然探索从基于已知标准的长度测量概念开始.由于直线段是从一点到另一点的最短路线,所以两点之间的距离以连接它们的直线段之长度来衡量.为了度量从点 C 到直线 AB 的距离,人们简单地发现,这种最短的路线既不向 A 倾斜也不向 B 倾斜,从而应当是从 C 到 AB 的垂线段.然而为了确定这条垂线段,人们还需要有构造垂线的方法.利用圆规,克莱罗就此给出了一种方法.既然有了垂线的概念,克莱罗就能定义矩形为每边都与其相邻边垂直的四边形,而正方形是四边皆等的矩形.克莱罗还给出了平行线的一个自然定义,即它们之间的距离处处相等.

矩形的度量基于具有单位边长的正方形.克莱罗说明矩形的面积是其长与宽的乘积.由于三角

形通常是矩形的一半,所以三角形的面积是其底与高之乘积的一半.然而由于区域并非总是直边形,克莱罗因此指出,度量这样的区域,可以先用直线段近似其曲线边,再将其分割为三角形并逐个测量.人们能够给出足够的近似以使"感觉得到的所有误差都被消除."

克莱罗以其自认为自然的方式,揭示了欧几里得《原本》第一卷至第四卷、第六卷、第十一卷和第十二卷中的大部分重要结果.例如,对于已知边长的一个三角形,他展示了如何构造其全等三角形.这是完全有必要的,因为就 ABC 所在之地可能无法对其进行直接测量,如其底边上的垂线有可能会经过一个障碍物.构造全等三角形的过程本身很简单.克莱罗将底 AB 移至新位置 DE,然后借助圆规分别度量 AC 和 BC 之长,以确定点 F,从而使得 $DF = AC$, $EF = BC$(图 14.4).对于克莱罗而言,作出的三角形显然在各方面都与给定的三角形相等.类似地,为了说明三角形由两条已知边及其夹角所确定,他借助了一种工具 abc,它由可以绕点 b 转动的两个直尺 ba 和 bc 组成(图 14.5),从而以一种明显的方式展示了如何构造与给定角相等的角.于是,给定角 B 为已知的三角形 ABC 和长度与 BC 相等的新线段 EF,他将其工具的直尺 bc 沿 EF 放置,以给定角度 B 向 EF 画一条新线段 DE,并使 DE 等于 BA.如此,通过连接 DF 而完成的三角形 DEF 与原三角形全等.克莱罗利用此类手法构造给定三角形的相似三角形,以此来测量难以靠近的点之间的距离.

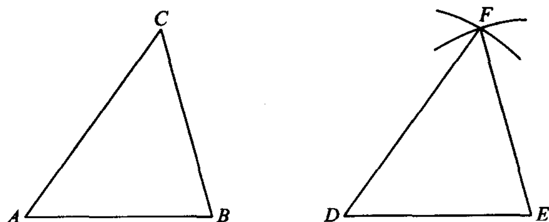


图 14.4 构造一个与给定三角形对应边长相等的三角形.

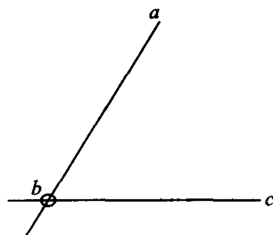


图 14.5 克莱罗用以画角的工具.

克莱罗指出,几何学家们并不希望像他所建议的那样通过近似的方法来度量曲线围成的区域面积.如果可能的话,直接度量这类区域将会显得更为"严格".圆是他以此方式处理的惟一图形.他在此证明了圆的面积是其周长与半径之一半的乘积.由于克莱罗不想使用古希腊的穷竭法及其相伴的归谬法,他打算使用圆是具有无穷多条边的多边形这一"事实".所以,他首先证明圆的任何内接正多边形的面积都等于其周长与其边心距之一半的乘积,然后指出,如果正多边形具有无穷多条边,那么其面积、周长和边心距将相应地变得与圆的面积、周长和半径相等.在其书关于立体几何的章节中,他同样认为一个正方锥由平行于底面的无穷多个薄片堆砌而成,并据此推证等底等高的两个锥体具有相同的体积.正如我们所见,这种基于不可分割量的论证可以回溯至数千年以前.然而,为了证明两个等高棱锥的体积之比与其底面积之比相同,克莱罗的确给出了一个更为严格的论证,并在事实上用到了阿基米德方法.基于这一结果,他以立方体的中心为顶点将其分割为六个相等的四棱锥,从而推知,高为 h、底面积为 B 的棱锥体积公式为 $V = (1/3)hB$.随之,将半径为 r 的球体看作由无穷多个高度为 r 的棱锥所组成,即可以计算出球的体积.由于这些棱锥的底面积之和即是球体的表面积,所以,要求的球体积等于半径与该表面积乘积的三分之一.为了导出该表面积等于四倍的大圆面积,克莱罗的论证涉及到了一种无穷小锥面,其面积已经通过其他无穷小论证而获得.

14.3.2 萨凯里与平行公设

在 18 世纪,根据欧几里得的其他公理和公设来"严格"推证其平行公设,以说明欧几里得毫无

必要提出那条并非不证自明的第五公设，重新激起了数学家们的兴趣．萨凯里(Girolamo Saccheri，1667—1733)和朗伯特(Johann Lambert)就是其中两位．

萨凯里于 1685 年加入了耶稣会会士的行列，随之相继在热那亚、米兰和都灵教授哲学，此后就是米兰附近的帕维亚大学，他在那里一直保持着数学教授的职位直至去世．1697 年，他发表了一部逻辑学著作，其中研究了某些类型的假定推理(false reasoning)，即从彼此矛盾的假设开始的推理．这最终引导他开始思考欧几里得公设，并着手研究欧几里得平行公设的替代假设是否与其他公理和公设相兼容．他于 1733 年发表的《欧几里得无懈可击》(Euclides ab omni naevo vindicatus)正是这一研究工作．萨凯里在该书的首章论述了平行公设的"瑕疵"，在次章研究了他所认为的另两块瑕疵，其中一个涉及第四项比例的存在性问题，另一个涉及复比问题．

我们在此仅考虑萨凯里著作的第一章．该章的目标旨在"清晰地证明这条颇存争议的欧几里得公理"[32]，其方法是假定它为不真，然后符合逻辑地推导出其正确性．萨凯里开始考虑四边形 $ABCD$，其 CA 和 DB 两边相等，且同时垂直于底边 AB．大约 600 年前花拉子米就曾考虑过这种四边形(图 14.6)．无需平行公设，而仅仅依靠欧几里得的其他命题，萨凯里很容易证明角 C 和角 D 相等．这两个角具有三种可能性，即同为直角，同为钝角，同为锐角．萨凯里将它们对应地称为直角假设，钝角假设和锐角假设，并随之证明了这些假设分别等价于如下命题，即线段 CD 等于、小于或大于线段 AB．如同对于早期研究这一问题的所有学者们一样，对于萨凯里而言，惟一"真实"的可能性"显然"是直角假设，这是因为平行公设在事实上蕴含着直角假设的正确性．另两个假设则在假定的基础上显示出平行公设不真．萨凯里试图仅用欧几里得"不证自明"的公理，从这两个"不真"的假设中反推出平行公设．如此以来，这两种可能性都将导致矛盾，从而所谓多余公设的"瑕疵"将在欧几里得的著作中被彻底消除．

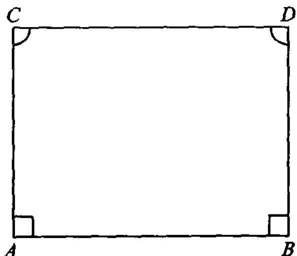


图 14.6　萨凯里四边形．

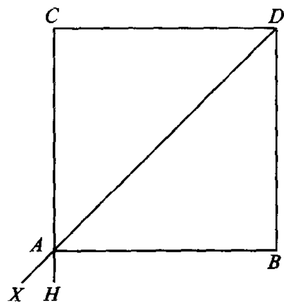


图 14.7　萨凯里的命题 VIII 和 IX．

萨凯里证明了，如果其中一个假设对于一个四边形为真，那么它对于所有的四边形都为真．然后继续研究如下命题：

命题 VIII　任给三角形 ABD，角 B 为直角；延长 DA 到任何一点 X，通过点 A 作直线 HAC 垂直于 AB，点 H 位于角 XAB 之中．可以断定，依照于直角假设、钝角假设或锐角假设为真的情形，外角 XAH 将分别等于、小于或大于相应的内角 ADB，反之亦然．

萨凯里的证明用到了欧几里得《原本》第一卷中的各种命题．他开始假定 AC 等于 BD，并将 CD 连结起来，这便形成了一个萨凯里四边形 $ABCD$(图 14.7)．根据直角假设 $CD = AB$．从而 $\angle ADB = \angle DAC = \angle XAH$，第一种情形获证．在钝角假设下，$CD < AB$，那么 $\angle XAH = \angle DAC < \angle ADB$，第二种情形获证．类似地，在锐角假设下，$CD > AB$，关于角的结论也随之成立．反之的证明相当简洁

明了.从该命题可以导出一个更为重要的命题:

命题 IX　在任何直角三角形中,另两个锐角之和在直角假设下等于一个直角,在钝角假设下大于一个直角,但在锐角假设下小于一个直角.

由于在三种假设的每种情形下,角 *XAH* 与角 *HAD* 之和都等于二倍直角,而角 *HAB* 是直角,所以角 *XAH* 与角 *DAB* 之和为一个直角.该结论可以直接从命题 VIII 得到.遗憾的是,这条定理有一点小问题未能引起萨凯里的注意.他的定理称三角形的两个非直角都是锐角.而事实上,它遵从于《原本》的命题 I-17,大意指三角形的任何两个角之和小于二倍的直角.正如第 2 章指出的那样,该定理依赖于欧几里得使用过,但却从未明确提及的一条假设,即一条直线可以任意延长,而该假设在钝角假设下并不成立.

尽管萨凯里没有觉察到任意延长直线的问题,但他后来的确在某种程度上证明了钝角假设将导致与《原本》I-17 相矛盾的结论.首先来看,无论是在直角假设(命题 XI) 还是钝角假设(命题 XII)的情形下,如果直线 *AP* 与 *PL* 交为直角,而与 *AD* 交为锐角,那么 *AD* 最终将与 *PL* 相交(图 14.8). 为了证明这一结论,他在 *AD* 上选取点 $M_1, M_2, M_3, \cdots$ 并使 $AM_1 = M_1M_2 = M_2M_3 = \cdots$.可以证明,如果对于每个 i 值,N_i 是点 M_i 在 *AP* 上的垂足,那么 $AN_1 \leqslant N_1N_2 \leqslant N_2N_3 \leqslant \cdots$.所以,有些 N_i 将落入 *P* 点之外,从而 *AD* 在某两点 M_{i-1} 与 M_i 之间与 *PL* 相交.如此以来,萨凯里可以基于如下一些假设来证明欧几里得平行公设:

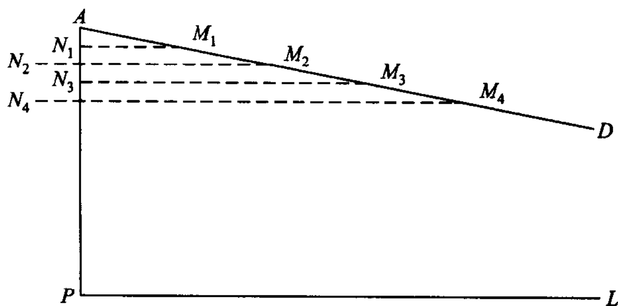


图 14.8　萨凯里的命题 XI 与 XII: *AD* 与 *PL* 将最终相交.

命题 XIII　如果直线 *XA*(具有任意给定的长度) 与直线 *AD* 和 *XL* 相交,并与其在同一侧交成内角 *XAD* 和 *AXL*,其和小于两直角(图 14.9). 可以断定,如果直角假设或钝角假设成立,即使这两个内角都非直角,后两条直线也会在该侧交于某点,交点定在有限处.

对于该命题的证明又一次用到了欧几里得的命题I-17.由于这两个内角之一,如 *AXL*,是锐角,人们可以向 *XL* 引垂线 *AP*.根据欧几里得的那条命题,该垂线的垂足将落在锐角 *AXL* 的一侧.由于在每种假设下,两个锐角 *PAX* 和 *PXA* 之和都不小于一个直角,如果将它们从给定角 *XAD* 与 *AXL* 的总和中减去,余角 *DAP* 将小于一个直角.命题 XI 与 XII 随之允许萨凯里得出结论称这两条直线将相交.

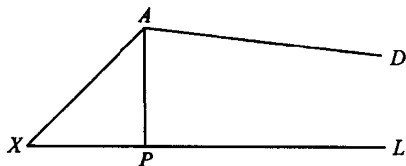


图 14.9　萨凯里的命题 XIII: *AD* 与 *PL* 将最终相交.

然而现在,由于处于钝角假设之下,三角形 *APX* 中的两个锐角之和大于一个直角,所以,萨凯里能够选择一个锐角 *PAD*,它与这两个锐角之和等于两直角.根据命题 XII,直线 *AD* 将最终与 *XP* 的延长线 *XL* 相交.据此,三角形 *XAL* 的两个角之和等于两直角,这与欧几里得的命题 I-17 相矛盾.

当然，由于平行公设已获证明，所以萨凯里能够像欧几里得在《原本》第一卷中所做的那样，通过命题 IX 中钝角假设自身的矛盾，证明任何三角形的三内角之和等于两直角．他因此提出：

命题 XIV　由于自身的矛盾，钝角假设是完全错误的．

萨凯里接下来证明了直角假设，钝角假设和锐角假设分别等价于任何三角形的内角和等于，大于或小于两直角，还等价于四边形的内角和等于，大于或小于四倍直角．随之，他还继续更为详细地研究了锐角假设下的逻辑结果．然而，他在此并未能推导出平行公设．但他却得到了其他一些新奇有趣的结论．例如：

命题 XVII　如果直线 *AH* 与长度不管多小的直线 *AB* 交为直角，可以断言，在锐角假设下，任何与 *AB* 交为锐角的直线 *BD* 最终将与 *AH* 的延长线相交的结论不可能为真(图 14.10)．

假设 *BM* 也垂直于 *AB*．从 *M* 向 *AH* 引垂线，垂足为点 *H*．由于四边形内角和小于四倍的直角，所以，角 *BMH* 是锐角．同理，如果 *BX* 是从点 *B* 所引的 *HM* 的垂线，垂足为 *D*．那么角 *XBA* 也是锐角．然而，由于 *H* 和 *D* 都是垂足，所以 *BD* 的延长线不可能与 *AH* 的延长线相交，否则将与《原本》的命题 I-17 相矛盾．

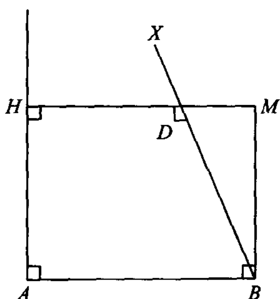


图 14.10　萨凯里的命题 XVII：*BD* 与 *AH* 不会最终相交．

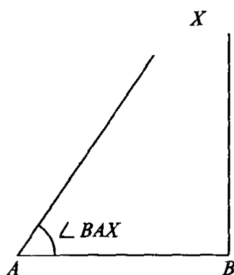


图 14.11　萨凯里的"平行角"：*AX* 与 *BX* 仅交于无穷远点 *X*．

由于命题 XVII 暗示着平面上有两条永不相交的直线，所以，萨凯里能够在命题 XXIII 中证明，对于这样的直线，他们或者具有相同的公垂线，或者"彼此向对方无限靠近"．[33] 此外，对于后一类情形，两条直线之间的距离会变得小于给定的任何长度；也就是说，它们互为渐近线．萨凯里因此能够在命题 XXXII 中证明，给定一条直线 *BX* 垂直于线段 *AB*，存在某一锐角 *BAX*，使得直线 *AX*"只在无穷远点与 *BX* 相交"[34]，而与 *BA* 交于更小锐角的直线将与 *BX* 相交，与 *BA* 交于更大夹角的直线都与 *BX* 有着一条相同的公垂线(图 14.11)．萨凯里随后得出结论称：

命题 XXXIII　由于与直线的自然特性相矛盾，锐角假设是完全错误的．

萨凯里对此几乎没有给出"证明"．看得出来，由于对平行公设的正确性深信不疑，所以，他实际上就此结束了自己的探索．他仅仅写到，对于锐角假设，必然存在两条直线，它们将最终"合而为一，而且在其所在平面上，它们在同一个无穷远点具有一条公垂线．"[35] 但是，他显然再次思考了这一问题，并利用此后 30 页的篇幅来竭力深化自己的推理过程．沿着两条直线不可能围成一块空间的思路，他证明了两条直线不可能有公共线段，以及给定直线在给定点只存在惟一的一条垂线．所有这些想法涉及的都是有限直线，并且没有一个与他的那两条合而为一且在无穷远点有公垂线的直线有关．不过，萨凯里确信自己已经达到了最初的目的．

14.3.3 朗贝特与平行公设

在对萨凯里的工作概要有一个基本的了解之后,朗贝特打算对其进行改进.但是,他于 1766 年完成的论述平行公设的著作《平行线理论》(Theorie der Parallellinien) 却从未发表.这或许是由于朗贝特对自己的结论并不满意.在这本书中,他仔细考虑了具有三个直角的四边形,并针对第四个角的特性作出了三种假设.它们在本质上与萨凯里的假设相同,即直角假设、钝角假设和锐角假设.同样地利用直线可以是任意长的原则,朗贝特丢弃了第二个假设.但他在否定第三个假设时却遇到了很大的困难.正如他所言,"这条假设一点也不自相矛盾."[36]

人物小传

约翰·朗贝特(1728—1777)(Johann Lambert)

朗贝特是一位自学成才的数学家和哲学家.在阿尔萨斯帮助父亲做裁缝的时候,他就掌握了数学知识.1748 年,他作为一位有钱人的家庭教师移居瑞士,后来带领自己的学生游历了欧洲.在此期间,朗贝特能够在这位富人家的家庭图书馆内学习工作,并进行理论和实践上的调查研究.然而,他从未抱有世俗的名利观.后来,有人建议位于柏林的普鲁士科学院提供一个职位给他.当他于 1764 年初到达那里的时候,曾受到欧拉的欢迎.但是,他奇特的外表与举止却使其任命被推迟了一年.最终,他克服了弗雷德里克二世初始时的敌意,并在其 49 岁英年早逝之前,创作了 150 多篇论著.

像萨凯里一样,朗贝特开始从这条假设推导各种结果.其中最令人称奇的是,在他的基本四边形中,内角和与 360° 的差有赖于四边形的面积;事实上,四边形越大,其内角和越小.如考虑四边形 $ABCD$,其角 A, B 和 C 是直角,角 D 是以 β 为大小的锐角(图 14.12).过 A 与 B 之间的 E 点作 AB 的垂线 EF,从而角(CFE 也是锐角,设其大小为 α.那么角 $\angle EFD$ 的大小为 $180° - \alpha$.然而,四边形 $EBFD$ 的内角和小于 360°,从而有 $90 + 90 + 180 - \alpha + \beta < 360$,或者 $\beta < \alpha$.所以,恰如所述,四边形 $ABCD$ 的内角和小于四边形 $AECF$ 的内角和.

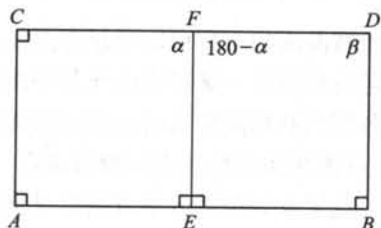


图 14.12 朗贝特对于四边形增大时其内角和减小的证明.

朗贝特据此得出结论称,"如果第三假设成立,每条线,每个面,每个体都将拥有一个绝对度量."[37] 换句话说,如果假定四边形 $AEFC$ 的两条边 $AE = AC$,那么,角 $\angle EFC$ 将是一个确定的锐角,一个不会在其他类似四边形中出现的角.如此一来,角 $\angle EFC$ 的大小 α 就可以被视作四边形的绝对度量.朗贝特未能提炼出这种绝对度量,他还不能算出在 $AE = AC = 1$ 英尺时角 $\angle EFC$ 的大小,但他的确已经意识到这条假设将彻底摧毁有关相似形的整个概念.他还能够证明,三角形内角和与 180° 之间的差即三角形的亏量(defect) 与三角形的面积成比例.朗贝特意识到在第二种假设情形下,会有一个类似的结果也为真,只是需要用内角和超出 180° 的盈余来代替亏量.他还清楚,球面三角形具有同样的性质,它们的内角和大于 180°,其盈余与其面积成比例.于是,他类推道:"我因此应当提议称第三条假设在一种虚球面上成立."[38]

朗贝特似乎深信欧几里得的几何学是空间真理,即使如此,当他一旦感到自己难以成功地反驳锐角假设,也便放弃了关于欧几里得平行公设的研究.不过,他认为由于钝角假设下的几何学对应着球面几何学,那么虚半径的球面将对于锐角假设起到同样的作用.虽然截止到 1770 年,他已经在

一定意义下引入了双曲函数作为三角函数的复对应,即 $\cosh ix = \cos x$, $\sinh ix = i \sin x$,但他既不能使用这些函数来发展基于锐角假设下的虚球上的几何学,也不能在三维空间中给出这种虚球的一个构造.只有等到 19 世纪早期,当这种类型的分析学被引入并能够为平行公设提供相关替换时,今天所谓的非欧几何学才得以发展.有关过程将留待第 17 章讲述.

14.3.4 克莱罗与空间曲线

18 世纪几何学最为主要的推进在于它与分析学取得了联系.欧拉《无穷小分析导论》的第二卷为平面曲线理论提供了清晰的组织结构.欧拉以对二次和三次曲线的分类为开始,随之论述了 146 种不同形式的四次曲线.接下来,他在未使用微积分学的情形下,大体上讨论了曲线的各种性质,其中包括渐近线,曲率和奇异点.他的书中甚至有一章专门论述超越曲线,他在其中探讨了由类似于方程 $2y = x^i + x^{-i}$(或 $y = \cos(\ln x)$)和 $y = x^x$ 所确定的曲线,并绘制了 $y = \arcsin x$ 的最早的草图.对于某些曲线,如阿基米德螺线,他还使用了极坐标对其进行了现代方式的描述.如此,若 s 表示极角,z 表示极径,那么螺线的方程是 $z = as$.类似地,方程 $z = ae^{s/n}$ 代表了对数螺线,欧拉也给出了它的图形.

最早发表的关于空间曲线的著作是 1731 年克莱罗的《双重曲率曲线研究》(Recherches sur les courbes á double courbure).对于克莱罗而言,一条空间曲线只能被定义为某些曲面的交线.因此,对各种情形的简单曲面的研究成为他工作的开始.他从几何的定义出发,说明了球面的方程为 $x^2 + y^2 + z^2 = a^2$,抛物面的方程为 $y^2 + z^2 = ax$,而且,一般来讲,由曲线 $f(x, u) = k$ 绕 x- 轴旋转生成的旋转曲面的方程可以通过在曲线方程中用 $\sqrt{y^2 + z^2}$ 替代 u 得到.他发展了锥面的一般概念,即它是将平面上一条任意曲线与平面外一点连接起来后所形成的曲面,并且证明了每个三元方程当其各项具有相同次数时必为圆锥曲面.他还证明了如下普遍结果,即一个三元方程通常定义一个曲面,且该曲面的特性由方程所决定.

克莱罗应用微分学方法以寻求空间曲线的切线和垂线,并因此研究了由"无穷多条极小线段"构成的曲线.[39] 确定一条曲线在点 N 的切线,就是要确定 Nn 的延长线与 xy- 平面的交点 t,其中 Nn 是 N 与其在曲线上的无限近点 n 的连线.或者,如果 M 是 N 在平面上的投影,确定曲线在点 N 的切线,也就是要确定线段 Mt 的长度(图 14.13).当然,寻找次切距 Mt 的思想是 17 世纪确定平面曲线

图 14.13 克莱罗与空间曲线的切线.

之切线的标准方法的类推.尽管在三维情形下,为了保证所有相关的直线位于同一平面上,计算过程稍微有点复杂,但其结果仍是二维情形下的直接类推.克莱罗将曲线上的无穷小线段 Nn 在 xy-平面上的投影记为 Mm,并将其延伸至交点 t.那么,三角形 NtM 定义了一个平面,要求的切线就位于其中.克莱罗只用到一条坐标轴,即 x-轴.因此,如果 AP 被用以表示 N 点的 x-坐标,那么,N 点的 z-坐标和 y-坐标分别为 N 到 xy-平面的垂线段 MN,和投影点 M 到坐标轴的距离 MP.若 Ap,nm,pm 是 n 点的三个对应坐标,此外,若 Nh 平行于 Mm,且 MH 平行于 Ap,那么 Pp 代表着 $\mathrm{d}x$,nh 代表着 $\mathrm{d}z$,mH 代表着 $\mathrm{d}y$,而可以用 Mm 表示.由于三角形 nNh 和三角形 NMt 相似,克莱罗得到比例关系 $nh : Nh = MN : Mt$.又因为 $Nh = Mm$,所以

$$\frac{\mathrm{d}z}{\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}} = \frac{z}{Mt} \quad \text{或者} \quad Mt = \frac{z\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}}{\mathrm{d}z}.$$

切线自身可因此由下式给出

$$Nt = \sqrt{MN^2+Mt^2} = \frac{z\sqrt{\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2}}{\mathrm{d}z}$$

此外,从曲线向 xz-平面引垂线 NO,它同时垂直于三角形 NtM 所确定的平面,且

$$NO = \frac{z\sqrt{\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2}}{\sqrt{dx^2+dy^2}}.$$

克莱罗针对这种计算给出几个例子,其中包括一条由两个抛物柱面 $ax = y^2$ 和 $by = z^2$ 所交成的空间曲线.在这种情形下,有 $a\mathrm{d}x = 2y\mathrm{d}y$ 和 $b\mathrm{d}y = 2z\mathrm{d}z$,所以

$$\mathrm{d}y = \frac{a\mathrm{d}x}{2\sqrt{ax}}, \quad \mathrm{d}z = \frac{b\mathrm{d}y}{2\sqrt{by}} = \frac{ab\mathrm{d}x}{4\sqrt[4]{b^2a^3x^3}}, \quad \text{且} \quad \sqrt{\mathrm{d}x^2+\mathrm{d}y^2} = \frac{\mathrm{d}x\sqrt{4x+a}}{4x}.$$

由于 $z = \sqrt[4]{ab^2x}$,故次切距 $Mt = \left(z\sqrt{\mathrm{d}x^2+\mathrm{d}y^2}\right)/\mathrm{d}z$ 可以表示为 $Mt = \sqrt{4x+a}$.同理,可以求出切线 Nt 为 $Nt = \sqrt{by+4x+a}$.对于垂线 NO,可以进行类似的计算.

14.3.5 欧拉与空间曲线和曲面

直到 1775 年,欧拉才开始研究空间曲线这一课题,这一次,他以弧长 s 为参数对空间曲线进行了表示.[40] 如此,一条曲线可以由三个方程 $x = x(s)$,$y = y(s)$,$z = z(s)$ 给出.欧拉发现,对每个方程取微分后可得 $\mathrm{d}x = p\mathrm{d}s$,$\mathrm{d}y = q\mathrm{d}s$,$\mathrm{d}z = r\mathrm{d}s$,据此可以推出 $p^2+q^2+r^2 = 1$.三个坐标函数关于弧长的导函数 p,q,r 分别是曲线单位切向量的三个分量.这些分量也被称为切线(或曲线本身)在指定点的方向余弦.为了定义曲线的曲率概念,欧拉利用了中心位于 $(x(s),y(s),z(s))$ 的单位球面.当参数分别取两个相邻值 s 和 $s+\mathrm{d}s$ 时,将"单位向量" (p,q,r) 的起点移至球心,如果其变化在单位球面上对应着弧差 $\mathrm{d}s'$,那么,该点的曲率 k 被定义为 $\left|\frac{\mathrm{d}s'}{\mathrm{d}s}\right|$,它度量了在任何点,曲线与球面上之大圆间的差别.由于向量 $\mathrm{d}s'$ 由下式给出:

$$\left(\frac{\mathrm{d}x}{\mathrm{d}s}(s+\mathrm{d}s) - \frac{\mathrm{d}x}{\mathrm{d}s}(s), \frac{\mathrm{d}y}{\mathrm{d}s}(s+\mathrm{d}s) - \frac{\mathrm{d}y}{\mathrm{d}s}(s), \frac{\mathrm{d}z}{\mathrm{d}s}(s+\mathrm{d}s) - \frac{\mathrm{d}z}{\mathrm{d}s}(s)\right) = \left(\frac{\mathrm{d}^2x}{\mathrm{d}s^2}\mathrm{d}s, \frac{\mathrm{d}^2y}{\mathrm{d}s^2}\mathrm{d}s, \frac{\mathrm{d}^2z}{\mathrm{d}s^2}\mathrm{d}s\right),$$

所以

$$\kappa = \left|\frac{\mathrm{d}s'}{\mathrm{d}s}\right| = \sqrt{\left(\frac{\mathrm{d}^2x}{\mathrm{d}s^2}\right)^2 + \left(\frac{\mathrm{d}^2y}{\mathrm{d}s^2}\right)^2 + \left(\frac{\mathrm{d}^2z}{\mathrm{d}s^2}\right)^2}.$$

欧拉接下来定义了曲率半径 ρ 即是曲率的倒数.因此有

$$\rho = \frac{ds^2}{\sqrt{(d^2x)^2 + (d^2y)^2 + (d^2z)^2}}.$$

虽然要一直等到 19 世纪才得以证明,但曲率作为空间曲线的两条本质特性之一已为人所知.另一个量是挠率,它度量着曲线偏离成为平面曲线的速率.如果曲率与挠率以沿曲线之弧长的函数形式给出,那么该曲线将完全被确定在它所处的空间位置上.

除了讨论空间曲线,欧拉还在其《无穷小分析导论》的第二卷以及若干年后的一篇重要论文中详细阐述了克莱罗的曲面论.在《无穷小分析导论》中,他对二次曲面理论进行了系统化整理.像克莱罗一样,欧拉只用到一个坐标平面,在其上只定义了一条坐标轴,而且借用空间一点到该平面的垂直距离来表示第三个坐标.但他的确说过,利用三个坐标平面并通过曲面在各种此类平面上的投影对其进行描述是完全可能的.他给出了三维空间中平面的方程 $\alpha x + \beta y + \gamma z = a$,但只将其系数的含义解释为该平面与 xy- 平面的夹角 θ 的余弦:$\cos\theta = \gamma/\sqrt{\alpha^2 + \beta^2 + \gamma^2}$.在关于二次曲面的讨论中,欧拉一开始就指出,一般的三元二次方程可以通过坐标变换简化为如下形式之一,即 $Ax^2 + By^2 + Cz^2 = a^2, Ax^2 + By^2 = Cz$ 或 $Ax^2 = By$.而其系数间的关系则决定了曲面的类型:椭球、椭圆或双曲抛物面、椭圆或双曲双曲面(如今对应地称之为单叶或双叶双曲面)、锥面以及抛物柱面.

由于《无穷小分析导论》是为学习微积分准备基础知识的一部著作,欧拉在该书中没有尝试着去论述类似于切平面和法平面的概念.而此类工作出现在他于 1760 年发表的一篇题为“曲面曲率研究”(Recherches sur la courbure des surfaces)的论文中.正是这篇论文被称为曲面微分几何的开端.[41] 欧拉在文中指出,尽管求平面曲线在给定点之曲率的方法已广为人知,但即使是给出空间曲面在某点的曲率的定义也非易事.通过给定点的不同平面切割曲面后将形成不同截线,即使严格限定切割平面与曲面保持垂直,它们的截线的曲率也可能会截然不同.在这篇论文中,欧拉对这些不同的曲率进行了计算,并在它们中间建立起一些关系.然而,他必须首先指明与曲面垂直的平面的特性,即它们都通过曲面在给定点 P 的法线.他证明,如果 $\beta\frac{\partial z}{\partial x} - \alpha\frac{\partial z}{\partial y} = 1$,则平面 $z = \alpha y - \beta x + \gamma$ 垂直于曲面 $z = f(x,y)$.欧拉将通过点 P 且同时垂直于曲面和 xy- 平面的平面定义为主平面,随之证明,如果曲面的一个给定的垂直平面与主平面的夹角为 φ,则通过该垂面截得的相应截线的曲率 $\kappa_\varphi = L + M\cos 2\varphi + N\sin 2\varphi$,其中 L, M, N 完全依赖于 z 在点 P 的偏导数.对此表达式关于 φ 求导后,欧拉发现,当 $-2M\sin 2\varphi + 2N\cos 2\varphi = 0$ 时,或当 $\tan 2\varphi = N/M$ 时,曲率取得最大值或最小值.但是,由于 $\tan(2\varphi + 180°) = \tan 2\varphi$,欧拉得出结论称,如果对于给定值 φ,曲率取得最大值,那么在 $\varphi + 90°$ 时,曲率取得最小值.最后,欧拉已经能够证明,如果 κ_1 是最大曲率,κ_2 是最小曲率,而且最小曲率发生在主平面内,那么,曲面的任何与主平面相交于角 φ 的垂直平面切割曲面所得截线的曲率为

$$\kappa = \frac{1}{2}(\kappa_1 + \kappa_2) - \frac{1}{2}(\kappa_1 - \kappa_2)\cos 2\varphi.$$

14.3.6　蒙日的工作

蒙日(Gaspard Monge,1746—1818)对于分析学和微分几何学进行了系统化总结,并为之引入了许多新材料.他的这一工作开始于 1771 年的若干论文,并在最终于世纪末汇结成他为巴黎综合工科学校的学生们所写的两部专著.例如,在 1784 年发表的一篇论文中,蒙日首次给出了直线的点斜式方程:“如果有人想表述以下事实,即这条直线[具有斜截式方程 $y = ax + b$]通过以(x', y')为坐标的点 M,就需确定截距 b,将方程变为 $y - y' = a(x - x')$,其中 a 是直线与 x 轴夹角的正切

值."[42]另一方面,蒙日在1799年的著作《画法几何学》对于代数学毫无涉及,而是紧紧地依赖于纯几何学思想.蒙日概述了在二维平面上绘制三维物体的许多方法技巧.系统地利用投影和其他空间变换,在二维平面上绘制空间图形的各个不同方面.详细描述了诸如曲面的切平面、两个曲面的交线、可展曲面(可被展开成为一个平面而且不会失真的曲面)以及曲面的曲率等概念.

人物小传

伽斯帕·蒙日(1746—1818)(Gaspard Monge)

蒙日(图14.14)出生于博纳,这是一座位于巴黎东南150英里的小镇.在里昂的时候,他是一位富有才气的学生.在为家乡绘制出规划蓝图之后,他被招致梅济耶尔的皇家工程学院(Royal Engineering School at Mézières).在那里,他很快赢得一次尽展自己才华的机会,应邀为一类特殊的防御工事绘制工程图纸.他没有使用传统的复杂方法,而是采用了一种新的绘图方法,这种方法后来被他扩展成为一个学科,即画法几何.他因此被提拔到教学岗位,从而影响了法国军事工程师的科学培训.1780年,他被选为法国巴黎科学院院士.在此后的35年间,他还先后在皇室、革命政府和帝国政府担任了许多重要职务.



图14.14 法国邮票上的蒙日.

他的另一本著作《分析在几何学中的应用》(Application de l'analyse ú la géométrie)发表于1807年,根据1795年以来的讲义整理而成.在这部书中,蒙日展示了如何将分析学应用于几何学.该书的首章只用到代数方法,对于二维和三维空间中的直线以及三维空间中的平面的最初的解析几何表述进行了详细介绍.蒙日指出,空间的点可以通过考虑它向三个坐标平面所引的垂线段得以确定.空间的直线可由其在三个坐标平面中任何两个上面的投影得到确定.例如,它在 xy- 平面上的投影线的方程可由斜截式或点斜式给出.蒙日展示了如何求两条直线的交点,如何求过给定点且与给定直线平行的直线,以及如何求经过两个给定点的直线.他还指出,当 $aa'=-1$ 时,平面中的直线 $y=ax+\alpha$ 和 $y=a'x+\alpha'$ 彼此垂直.

蒙日对于平面的方程给出两种记法,其一为 $z=ax+by+c$,其中 a 和 b 分别是该平面与 xz- 坐标面和 yz- 坐标面的交线的斜率;其一为对称式 $Ax+By+Cz+D=0$,其中系数 A,B 和 C 决定着该平面与各坐标面夹角的方向余弦.随后,他继续探讨有关点、线与面的所有常见问题,如求经过给定点的平面法线,求两条直线间的最短距离,以及求两条直线或直线与平面间的夹角等.

蒙日著作的第二章专注于曲面研究.在这里,他借用微积分的全部工具分析地发展了他在《画法几何学》中曾考虑过的所有主题.如此,他详细研究了如何根据各种类型的绘制方法来确定某给定曲面的偏微分方程,以及如何在某种情形下通过积分求解这一方程.为了得到曲面的切平面和法线,蒙日一开始就指出,在点 (x',y',z') 的附近,曲面 $z=f(x,y)$ 的微分方程为

$$\mathrm{d}z=\frac{\partial z}{\partial x}\mathrm{d}x+\frac{\partial z}{\partial y}\mathrm{d}y.$$

上式中的偏导数在 x' 和 y' 处取值.另一方面,通过点 (x',y',z') 的任何平面的方程可以表示为 $A(x-x')+B(y-y')+C(z-z')=0$.由于该平面是一个切平面,其上与给定点无穷近的点必然也在曲面上,也就是说,必然满足曲面的微分方程.因此,令 $x-x'$ 为 $\mathrm{d}x$,$y-y'$ 为 $\mathrm{d}y$,$z-z'$ 为 $\mathrm{d}z$,蒙日指出,方程 $A\mathrm{d}x+B\mathrm{d}y+C\mathrm{d}z=0$ 必然与 $\mathrm{d}z=\frac{\partial z}{\partial x}\mathrm{d}x+\frac{\partial z}{\partial y}\mathrm{d}y$ 等价.所以 $A/C=-\partial z/\partial x$,$B/C=-\partial z/\partial y$,从而,切平面的方程为

$$z-z'=(x-x')\frac{\partial z}{\partial x}+(y-y')\frac{\partial z}{\partial y}.$$

曲面的法线方程，即是切平面的法线方程，随之可以被求出为

$$x - x' + (z - z')\frac{\partial z}{\partial x} = 0; \quad y - y' + (z - z')\frac{\partial z}{\partial y} = 0.$$

蒙日将微分方程与空间几何学联系起来的基本思想，流经岁月，影响久远．然而，更为重要的或许还在于他作为巴黎综合工科学校的老师影响了整整一代法国工程师、数学家和科学家．有关史实留待14.4 节再续．

14.3.7　欧拉与拓扑学的发端

大约在18世纪30年代中期，欧拉得知了从东普鲁士（现在俄罗斯）的哥尼斯堡小镇传出的一个小问题．在流经小镇的普利格尔河的中间有两座小岛，七座桥将它们与河的两岸相连．小镇上的居民们提出一个问题，是否能够设计出一次步行路线，以从每座桥上只经过一次．像通常一样，欧拉没有孤立地看待这一问题，而是向一般性问题发起了冲击，并解决了在区域与桥的数目为任意数时类似路线的存在性问题．在 1736 年发表的一篇论文中，他首先指出，如果将各区域标记为字母 $A, B, C, D, \cdots$，那么路线可以被标记为一系列字母，以代表相继途经的各区域（图 14.15）．如此以来，*ABDA* 将代表一条路线，它从 A 区域出发，经 B 区域，到达 D 区域，再返回 A 区域，而无需考虑途经了哪座具体的桥．显然，满足条件的一条完整路线包含的字母数要比桥的数目多一个．在哥尼斯堡问题中，表示路线的字母数应该为 8.

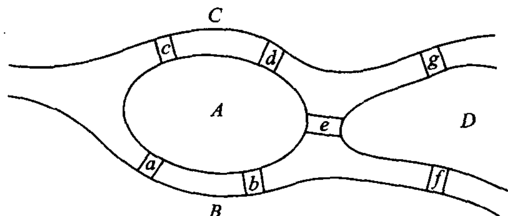


图 14.15　哥尼斯堡七桥．

接下来，欧拉认识到，如果通向某给定区域的桥的数目 k 为奇数，那么代表该区域的字母必然出现 $(k+1)/2$ 次．例如，若只有一座桥通向区域 A，那么 A 将只出现一次；如果有三座桥，那么 A 就会出现二次；如此等等．路线是否从区域 A 或其他区域开始，对此都无妨碍．另一方面，如果 k 是偶数，那么，当路线从其他区域开始时，代表该区域的字母将出现 $k/2$ 次；当路线从该区域内开始时，区域的代表字母将出现 $k/2+1$ 次．例如，若有四座桥通向区域 A，那么当路线以 A 之外某区域为起始时，A 将出现两次；当其以 A 为出发点时，A 将出现三次．根据具体路线中区域的代表字母出现次数的两种不同计算方法，欧拉可以确定出一条路线是否可能仅有一次经过每座桥：

> 若［以上计算出的各区域的］出现次数的总和等于桥的总数加一，则要求的行程是可能的，而且出发点必须位于一个与奇数座桥相连的区域．但是，若此总数比桥的总数加一少一个，那么行程亦然可能，但需从另一与偶数座桥相连的区域出发，只有如此，字母出现的次数才能随之增加 1 次．[43]

在哥尼斯堡问题中，通向每块区域 A, B, C, D 的桥的数目都为奇数，即分别是 5,3,3 和 3．与之相应的途经次数 3,2,2 和 2 之和为 9．该数字大于“桥的总数加一”，所以要求的路线是不可能存在的．欧拉指出，一般来说，当连通着奇数座桥的区域多于两块时，这种路线将始终不可能存在．如果这样的区域正好有两块，那么只要以其中之一为起点，就有可能找到一条满足条件的路线．最后，如

果通向各区域的桥的数目都是偶数，那么恰好一次途经每座桥的路线总有可能存在.由于一旦得知路线存在与否的可能性，实际的构造将是简单明了的，所以，欧拉已经彻底解决了他自己设立的问题.



图 14.16 德意志民主共和国邮票上的欧拉与多面体公式.

欧拉还解决了另一个问题，即给定任意简单多面体，它具有 V 个顶点，E 条边和 F 块面，那么 $V - E + F = 2$(图 14.16).该问题与前一特殊问题在 18 世纪只是一些孤立的事实，然而，欧拉的确强调过，它们显然属于几何学的某一分支，其中对象间的关系只与其相对位置有关，而与其数量大小毫无关系.但是，只有等到 19 世纪末和 20 世纪初，这些问题以及其他事实才得到系统化研究并最终发展成为一个学科，即拓扑学.

14.4 法国大革命与数学教育

18 世纪的主要数学家们都与大学无缘，而是与各国君主们所设立的专科院校联系在一起.这些君主们设立院校的目的旨在为自己的国家赢得声望，并为国家的发展在军事和民用工程方面提供必要和充足的科学贮备.一般而言，大学并不提供针对数学的高等教育，即使是在 18 世纪，它们基本上仍是哲学家们一统天地.尤其是在法国，从 14 世纪以后，任何一流的数学家都与巴黎大学 (University of Paris) 毫无关系.提供数学与科学教育的学校只有军事院校，这类学校的一个主要功能是培养军事工程师.因此，蒙日的早期职业是在梅济耶尔(Mézières) 的军事学院任教.在那里，他因绘制军事防御工程设计图而形成了自己关于画法几何学的初步思想.同样，拉普拉斯和勒让德也曾一度在巴黎军事学院(École Militaire in Paris) 执教.

由于军事院校以及各所大学在法国大革命期间是保皇派支持者们的活动中心，所以，截止 1794 年当革命发展达到高潮的时候，它们中的许多已经被相继关闭.不过，随着邻国军队对法国的攻击，以及受过良好教育的国民从法国的出走，国家完全有必要设立学院并将那些缺少贵族背景但却"一直热爱自由与平等并痛恨专制统治"[44] 的学生们培养为工程师或科学家，使其为军队或地方建设服务.正是出于这一目的，法国国家议会于 1794 年 9 月 28 日宣布创办大众联合工程学院(École Centrale de Travaux Publiques)，不久更名为综合工科学校(École Polytechnique).然而，该学院并不仅限于是一所工程学院，它的宗旨还包括培养高素质的国民，特别是激励天才们来发展科学.

法国大革命前曾在海军学院帮助从事科学教学改革的蒙日，被指命代为负责组建综合工科学校.他因此而忙于开设"革命时期的课程"(revolutionary course)，提供新生们从十二月开始学习的、将历时三个月的自然科学概观，这对学生们在此后两年或三年时间里将要从事的研究课程将是一个预先展示.学生们将要学习四门基本学科，即画法几何学、化学、分析与力学、物理学，此外还有一门工程制图课.后一门课程每天晚上授课三学时；而前三门课程都只在每天上午有一学时的讲座时间，随后是一学时的辅导研究；然而，物理学每旬只有四学时的授课时间.(大革命时期的历法将一个月分为三旬，而不是将每七天记为一周.)

蒙日亲自讲授画法几何学，第一个月的课程实质上涉及到我们在本书第 14.3.6 节所描述的内容.如此，学生们按计划将学习投影的方法，曲线的切线和法平面的求法，曲面的切平面和法线的求法，曲面交线的构造，以及可展曲面的概念.他们还将研究这些思想在各种问题如建筑设计和图形绘制领域中的应用.课程在第二个月涵盖了建筑学与公共建设工程，在第三个月处理军事防御工程问题.分析学也由蒙日讲授，第一个月从求解直到四次的多项式方程开始，随之转向方程组的代数与几何解法，以及对这些方程所代表的曲线与曲面的研究.第二个月讲解级数论、指数和对数函数，

基本概率论和微分学及其几何应用. 最后一个月讲解积分学,包括长度、面积和体积的求法,以及微分方程的解.

这一课程纲要相当宏伟,然而不幸的是,它并没有能够得以实施. 当 1794 年 12 月 21 日院校开课之时,蒙日病倒了. 他的画法几何课因此被迫延期. 他的分析课程由费利(C. J. Ferry) 代为讲授,而格里费特 - 拉博美(C. Griffet-Labaume) 则为那些在空余时间仍想学习的学生们重复讲解. 遗憾的是,在开学头十天的课程结束之前,已经可以清楚地看到,大多数学生根本无法理解课程的内容. 当时学校董事会的拉格朗日很快便决定让拉博美开设一门代数基本课程以代替其复讲的分析学. 但即使这门新课程在第一个月只讲解到平面曲线的二元方程表示,也只有近三分之一的学生坚持了下来. 第二个月,随着代数基本课程的继续,增设了一门三角学课程,加之蒙日开始讲授他的画法几何课,情况似乎有所改观. 但是在计划与现实之间显然仍存在着巨大的差距. 学院在创办之初显得景况窘迫,自然有其原因所在. 其中之一即是由于食物短缺使得巴黎的冬日变得尤为严寒. 不过,最根本的原因却在于学生们缺乏足够的准备. 这些学生们入学前在各自的家乡主要只接受有关"政治合格" 的测试,但是这种政治性入学考试并不能弥补学生们对应该掌握的科学理论知识的缺乏.

尽管学院初创时举步维艰,蒙日仍和他人一起很快使得综合工科学校取得了巨大进展,并使之成为整个欧洲和美国所有工程学院的典范. 随着一所新师范学院在 1795 年的创办,国家教育水平逐渐得以确立. 综合工科学校向全国各省派遣主考官以确保被招入的新生具有良好的基础知识. 课程设置也在某种程度上变得更加务实. 除了为期三年的常规课程教学,"革命时期的课程" 已停止讲授. 最后,全法国最好的数学家们都来学院任教,其中包括拉格朗日、拉普拉斯和拉克鲁瓦,有些数学家还为学院编写了使用教材. 特别是拉克鲁瓦,他编写的课本包括算术、三角、解析几何、综合几何、微分学和积分学. 这些著作大多数都经历了诸多版本,并被译成多种语言. 事实上,拉克鲁瓦的微积分教材在 1816 年被译成英文,对于将欧洲大陆上的数学方法引入英国和美国产生过重要影响.

除了改变法国技术教育的自然状况,革命政府还使法国的度量衡制度得以标准化,从而引入了公制度量衡(图 14.17). 国民代表大会通过一项基本法案,要求从 1790 年 5 月开始对度量衡进行标准化. 法国科学院随之组成一个包括拉普拉斯、拉格朗日和蒙日在内的委员会来研究这一课题. 第一份建议是通过秒摆来定义长度单位. 但到 1791 年 3 月,委员会决定长度的标准单位应该定义为地球大圆四分之一弧长的十万分之一,如此显得比用时间定义更加"自然". 国民代表大会颁布法令重新测算通过巴黎的子午线,以准确地求出该长度单位. 一年后,在拉普拉斯的提议下,该长度单位被称为"米"(meter). 随着长度单位的确定,委员会决定所有的细分或倍增都应该采用十进制. 进一步,面积与体积的度量也基于长度度量来定义. 如此,面积的基本单位,即边长为 100 米的正方形的面积,被称为一"公亩"(are),同样,质量的基本单位 —— 克(gram),被定义为一立方厘米水在给定温度下的质量(图 14.18).



图 14.17　为纪念公制度量衡的引入而发行的法国邮票.

委员会的成员们还进行了进一步调研,通过黄金与白银的价值设计出与重量相关的货币十进制单位,并通过将圆周的四分之一弧划分为 100 等分,设计出角度的十进制单位,即百分度(grads). 最后,他们还制订了革命时期的历法,将十进制延伸到了历史学领域. 拉普拉斯提议并赞同将每月分为三旬,每年的最后五天定为假期. 即使已经意识到历法中的这种十进制所引发的问题将会多于它所能解决的问题,他还是给出了日与年之间的这种不可通约性规定. 特别有趣的是,虽然在接下来的一个世纪里,整个世界都在实际上接受了重量和长度的十进制,但关于角度与历法的十进制却

图 14.18　选自度量衡委员会的法文图书中的一块刻盘(1793 年),给出了重量的旧标准与公制标准之间的换算. 注意,重量的基本单位被称为 grave 或 nouveau poid, 而非千克(kilogram). (资料来源: 斯米森研究所图书馆(Smithsonian Institution Libraries)图片编号:89 - 8736.)

没能维持超过十二年.

拿破仑于 1799 年开始统治法国,并于 1806 年宣布在法国恢复使用格里高利历(Gregorian calendar). 在此其间,他赢得了法国一些重要科学家们的支持. 特别是蒙日,因曾陪同这位波拿巴于 1798 年远征埃及,从而成为他的有力支持者. 作为报答,蒙日被任命为终身参议员,稍后又被加封为荣誉之旅的大指挥和贝鲁斯伯爵. 勒让德、拉格朗日和拉普拉斯也获得了荣耀. 前者被封为帝国爵士,后两位则被封为伯爵. 然而,随着 1815 年拿破仑的倒台,蒙日失去了所有的职位,在一种智力流放中度过了生命中的最后三年. 拉格朗日已于 1813 年去世. 但拉普拉斯和勒让德却与复辟王朝言归于好,并得以继续从事自己的研究工作.

14.5　美洲的数学发展

18 世纪末期法国所发生的政治剧变在某种程度上得益于 1776 年美国独立战争的启发. 然而美国的数学却与法国的数学在水平上相差甚远. 到独立战争时期,在即将变为美国版图的领土上共有九所大学. 由于创办其中大多数学校的根本目的只在于培养神职人员,所以学校的许多数学教师本身就是牧师,他们主要对于神学而非数学感兴趣. 因此,数学教育水平远低于同时期欧洲大学相应的教学水平. 开设的普通课程中有算术,贯穿简单方程求解的代数,基本几何学,贯穿用对数解平面三角形的三角学,以及这些知识在测量学和天文学方面的应用. 到 18 世纪中期,至少在哈佛大学和耶鲁大学已有人着手开设关于流数的课程,并有人开始从事私人数学教育. 1727 年,格林伍德(Isaac Greenwood, 1702—1745) 开始教授流数术以及莱布尼兹微分学,此外还有航海学、测量学、力学、光学以及天文学. 他从 1728 年至 1738 年在哈佛大学任霍利斯数学教授. 1729 年,他编撰并发表了第一部美国人自己的算术著作. 而在通常情况下,美国人都使用从英国进口的教材.

在这个世纪，美国按照欧洲类似的协会模式创立了两个科学组织，即成立于 1743 年的美国哲学会和组建于 1780 年的美国文理科学院．但是，即使是后者的院刊对于当时美国人自己的数学成就也持悲观见解．在其《论文集》(Memoirs) 的第一卷中有一篇发刊词称，文集中发表的数学论文可能只会引起极少数人的兴趣，而且，无论如何，它们几乎不包括与研究有关的内容而只是实践性的．

虽然在 18 世纪的美国没有研究型的数学家，但却有一些著名的天文学家和测量学家，他们至少能够理解数学的重要性．温思罗普(John Winthrop，1714—1779) 曾在哈佛大学担任了 40 年的霍利斯数学教授，讲授具有牛顿传统的数学知识，涉及的主题包括“几何原理，比例学说，代数法则，二次曲线，平面与球面三角，测量位面与立体的普遍法则，球的应用，[以及] 根据托勒密、布拉埃和哥白尼的不同假说而对天体的运动与现象的计算．”[45] 黎特豪斯(David Rittenhouse，1732—1796) 是一位天文学家，钟表制造家和测量学家．他从事了大量详细的天文观测，并帮助确立了梅森 - 狄克逊线(Mason-Dixon line)，发表了部分数学论文，其中一篇讨论正弦函数乘幂的积分方法．班尼克尔(Benjamin Banneker，1731—1806) 是在科学界获得声望的第一位非洲裔美国人．靠自学掌握了足够的数学与天文学知识，并在 18 世纪 90 年代出版了一系列天文年历．

人物小传

本杰明・班尼克尔(1731—1806)(Benjamin Banneker)

班尼克尔的父亲是一位获释的奴隶，他的外祖母先时是一位英国契约女仆，后来嫁给了自己的黑奴，该黑奴原本可能是一位非洲酋长的儿子．班尼克尔在外祖母教导下练习读写，并由其安排在冬季学期去一个很小的乡村学校学习．尽管作为农民又是黑人，他在马里兰是不允许发展个人才干的，然而，他在很小的时候就展露出极大的技术天赋．在 22 岁那年，他利用木料自行制作了一台精确的时钟，该时钟一直运转着，直到他死后不久才毁于火灾．不过，班尼克尔有幸与埃利柯特家族为邻．这是一个商人和测量员之家．从这个家庭那里，班尼克尔能够借阅许多技术书籍还有部分科学仪器．在自学了数学、测量学以及天文学的基本原理之后，他应埃利柯特之邀协助其测绘哥伦比亚特区的边界线，该项任务是由华盛顿总统于 1791 年指派给埃利柯特的．完成任务之后，班尼克尔继续从事自己的研究，直到他能够编制出 1792 年的历书．该历书包括了太阳、月亮和行星每天所处的位置，太阳、月亮和某些亮星升起和降落的时间，以及地方潮汐表．班尼克尔的历书被证明相当流行，所以他能够每年出版一册类似的历书，直到 1797 年结束(图 14.19)．

还有两位重要的公众人物，他们在 18 世纪末期为推动数学研究的发展做出了不少贡献．富兰克林(Benjamin Franklin，1706—1790) 虽然不是数学家，但他通过发表评论以及自己在创建宾夕法尼亚大学方面的影响来促进数学研究．由于他在欧洲有很高的知名度，而且是皇家学会的成员，所以有能力在欧洲的科学团体与美国新生的相应组织之间开创相互交流．最后，杰斐逊(Thomas Jefferson，1743—1826) 曾在威廉和玛丽学院(William and Mary) 学过一些数学，但他对于这门学科的掌握在很大程度上还应归功于自学．在出任驻法大使期间，他重新唤醒了自己对数学及其应用的兴趣，在诸如测量学、天文学、球面三角学，以及公制度量衡等领域发表了许多文章．在 1799 年发自其家乡蒙蒂塞洛(Monticello) 的一封信中，他表达了自己对数学价值的看法：



图 14.19　美国邮票上的班尼克尔．

> 三角学 …… 对于每个人都颇有价值．人们为了日常的生活目的，几乎没有一天不用它．包括开平方和开立方，计算的科学也必不可少．直至二次方程的代数学以及对数方法

在一般情形下也相当有用. 然而, 超越这些之外的所有其他知识都仅仅是一种奢侈品, 一种确能给有些人带来享受和乐趣的奢侈品, 但那些为求生存而以此为业的人绝不应沉湎其中. 以我的观点来看, 圆锥截线, 高次曲线, 或许还有球面三角学, 二维以上的代数运算, 以及流数术都是如此.[46]

在 18 世纪末, 而且实际上在进入 19 世纪以后, 美国并不需要数学中的这些"奢侈品".

在美洲的其他地方, 受人关注的数学知识仍然是实用性数学. 例如, 在魁北克, 狄邦内坎普神父 (Father J. P. DeBonnécamps) 就讲授应用数学, 其中包括水文地理学, 测量学和天文学. 在格兰德河以南, 罗马教皇打算将这片土地划归西班牙和葡萄牙的决定, 引发了南美洲对资深测量员的需要. 美洲的第一部数学著作是 1560 年由神父第兹 (Father Juan Diez) 在墨西哥所写的《计算概要》(Sumario Compendioso de las Cuentas), 它广泛收录了其他财政事件中与金银兑换相关的数表, 以及与这些数表有关的算术问题和一些初等代数问题. 一般而言, 拉丁美洲其他地方的早期数学著述也都只包含非常实用的数学内容, 特别是符合军事目的的实用内容. 不过, 虽然拉丁美洲在殖民时期曾创建了一些大学, 但却很少讲授数学知识. 只有等到这片土地在 19 世纪实现独立之后, 数学才在那里获得发展的机会. 最终, 在格兰德河南北, 当学者们掌握了足够的数学基础之后, 欧洲的数学知识才得以这块新大陆取得进一步发展.

习　题

伯努利《猜度术》中的问题

1. 计算伯努利数 B_8, B_{10} 和 B_{12}. 伯努利数列常常如下设定: 令 $B_0 = 1, B_1 = -(1/2)$, 而 $B_k = 0$, 其中 k 是大于 1 的奇数.

2. 利用伯努利的方法, 明确写出整数 4 次、5 次和 10 次乘幂的前 n 项和公式. 然而证明前 1000 项正整数的 10 次幂和为

$$91,409,924,241,424,243,242,241,924,242,500.$$

伯努利称他 (未借助计算器) 算得此值仅用了"不到半刻钟" 时间.

3. 证明: 如果将伯努利数 B_i 定义为满足

$$\frac{x}{e^x - 1} = \sum_{i=0}^{\infty} \frac{B_i}{i!} x^i,$$

那么当 $i = 2,4,6,8,10,12$ 时, B_i 的值将与正文及练习题 1 中算得的结果相同.

4. 通过以下证明来完成伯努利关于大数定律例题的计算. 即当 $r = 30$ 且 $s = 20$ (从而 $t = 50$) 时, 若 $c = 1000$, 那么,

$$nt + \frac{rt(n-1)}{s+1} > mt + \frac{st(m-1)}{r+1},$$

其中 m, n 为整数, 且有

$$m \geqslant \frac{\log c(s-1)}{\log(r+1) - \log r} \quad 和 \quad n \geqslant \frac{\log c(r-1)}{\log(s+1) - \log s}.$$

由此推知在这种情况下必需的实验次数为 $N = 25\,550$.

5. 利用伯努利的公式证明, 如果在练习题 4 中需要有更大的确定性, 比如说 $c = 10\,000$, 那么必需的试验次数为 $N = 31\,258$.

6. 设 a 是一次试验中的成功率, 而 $b = 1 - a$ 是失败的概率. 如果重复三次试验, 证明成功次数 S 分别为 3,2,1,0 时,

对应的事件概率为 $P(S=3)=1a^3, P(S=2)=3a^2b, P(S=1)=3ab^2$ 和 $P(S=0)=1b^3$.

7. 将练习题 6 中的试验次数推广到 n 次. 证明成功次数为 r 的概率是

$$P(S=r)=\binom{n}{n-r}a^r b^{n-r}.$$

8. 利用练习题 7 的结论，在 $a=1/3, b=2/3$ 和 $n=10$ 时，计算 $P(4\leqslant S\leqslant 6)$.

棣莫弗的问题

9. 设一次试验成功的概率为 1/10，试问需要进行多少次试验才能保证至少成功一次的概率出现均等情形？利用棣莫弗的精确方法和近似方法分别对此进行计算.

10. 需要将三枚骰子一起抛掷多少次，才能保证其同时出现一点的事件至少发生一次的概率具有等可能性？

11. 在一次摸彩活动中，失败与成功的比率为 39:1，需要购买多少张彩票才能为赢得奖金带来均等的机会.

12. 推广棣莫弗在(其著作中的)问题 III 中的算法，以解决问题 IV：设在一次试验中某事件成功的概率为 a，而失败的概率为 b，试求需要多少次试验，才能使该事件发生两次的情形出现等可能概率.(提示：注意事件最多只成功一次的机会为 b^x+xab^{x-1} 次，而总次数为 $(a+b)^x$.) 当 $a:b=1:q$，其中 q 是一个相当大的数时，求问题的近似解，并证明 $x\approx 1.678q$.

13. 证明棣莫弗从比率 $\binom{n}{n/2}:2^n$ 中引出的标以 col.1，col.2 及 col.3 的和式可以明确表示为

$$\text{col.}1=\frac{s^2+s}{m}$$

$$\text{col.}2=\frac{\frac{1}{2}s^4+s^3+\frac{1}{2}s^2}{3m^3}$$

$$\text{col.}3=\frac{\frac{1}{3}s^6+s^5+\frac{5}{6}s^4-\frac{1}{6}s^2}{5m^5}.$$

计算 col.4 的对应值.

14. 将练习题 13 中各列的最高次项相加得

$$s\left(\frac{s}{m}+\frac{1}{2\cdot 3}\frac{s^3}{m^3}+\frac{1}{3\cdot 5}\frac{s^5}{m^5}+\frac{1}{4\cdot 7}\frac{s^7}{m^7}+\cdots\right),$$

令 $x=s/m$，则上式变为

$$s\left(\frac{2x}{1\cdot 2}+\frac{2x^3}{3\cdot 4}+\frac{2x^5}{5\cdot 6}+\frac{2x^7}{7\cdot 8}+\cdots\right).$$

证明括号中的级数可以用有限项表示为

$$\log\left(\frac{1+x}{1-x}\right)+\frac{1}{x}\log(1-x^2),$$

从而原级数可以表示为

$$mx\log\left(\frac{1+x}{1-x}\right)+m\log(1-x^2).$$

既然 $s=m-1$(或者 $mx=m-1$)，那么，证明练习题 13 中各列最高次项之和为

$$(m-1)\log\left(\frac{1+\frac{m-1}{m}}{1-\frac{m-1}{m}}\right)+m\log\left[\left(1+\frac{m-1}{m}\right)\left(1-\frac{m-1}{m}\right)\right],$$

该式即是 $(2m-1)\log(2m-1)-2m\log m$.

15. 练习题 13 中各列次高项之和为

$$\frac{s}{m}+\frac{s^3}{3m^3}+\frac{s^5}{5m^5}+\frac{s^7}{7m^7}+\cdots$$

由于 $s = m - 1$,证明该和式等于

$$\frac{1}{2}\log\left(\frac{1+\frac{s}{m}}{1-\frac{s}{m}}\right) \quad 或 \quad \frac{1}{2}\log(2m-1).$$

16. 推导棣莫弗的结论

$$\log\left(\frac{Q}{M}\right) \approx -\frac{2t^2}{n},或等价地 \log\left(\frac{M}{Q}\right) \approx \frac{2t^2}{n}.$$

(提示:以 m 除正文中所给表达式前两个对数项的自变量,然后简化并将其余对数项替换为其相应幂级数的前两项.)

贝叶斯的问题

17. 利用贝叶斯定理,直接计算 $P(r < x < s \mid X = n - 1)$. 特别地,假设你已经从所含黑白球比例未知的罐中抽得 10 个白球和 1 个黑球. 如果现在你猜测该未知比例大于 7/10,你猜对的概率有多大?
18. 证明:若概率未知的事件连续发生了 n 次,那么,它再次发生的机会超过 1/2 的可能性为 $2^{n+1} - 1$ 比 1.
19. 设想罐中有两个球,每个球可能是白色,也可能是黑色. 每次抽取一个球,并在下次抽取新球前将旧球放回. 假如前两次抽取的都是白球,那么第三次抽得白球的概率是多少?

关于年金保险与彩票抽奖问题

20. 棣莫弗在对死亡率统计表的研究中发现,人们至少可以对超过 10 岁的年龄为 k 岁的人假定,其再活 1,2,3,… 年的概率近似地等于 $\frac{n-1}{n}, \frac{n-2}{n}, \frac{n-3}{n}, \cdots$ 其中 $n = 86 - k$,表示棣莫弗所谓的生命余数(complement of life). 证明在利率为 r 的假设下,年龄为 k 岁的人每年 1 英镑的终身养老金到现在的总量可以由下式算出:

$$\frac{n-1}{n(1+r)} + \frac{n-2}{n(1+r)^2} + \frac{n-3}{n(1+r)^3} + \cdots.$$

21. 如果 $P = \sum_{m=1}^{n} \frac{1}{(1+r)^m}$ 是每年 1 英镑的养老金在 n 年后达到的总量,其中 r 表示利率. 基于练习题 20 的假设与符号,证明年龄为 k 岁的人现在的终身养老金为

$$A = \frac{1 - \frac{(1+r)P}{n}}{r}.$$

证明利率为 5% 时,对于年龄为 36 岁的人,每年 1 英镑的终身养老金现时的总量为 16.546 8. 然后计算年龄为 50 岁的人每年 1 英镑的终身养老金是多少.

22. 在 18 世纪后期的法国皇家彩票抽奖活动中,有五个标有数码的球被从 90 个球中随机抽出. 起先,摸彩者可以只买一个号码,或者一组二个号码,或者一组三个号码. 后来,摸彩者被允许赌买一组四个号码或五个号码以及一组按次序抽取的号码. 证明赌买单一号码,一对号码,和一组三个号码时失利的可能性分别为 17:1,399.5:1,和 11 747:1. 这些赌注的回报为 15 270 和 5500.

牛顿《通用算术》中的问题

23. 在三个工人中,A 可以在三周内完成一件给定的工作,B 可以在八周内完成三件,而 C 可以在十二周内完成五件. 那么他们共同完成这件工作需要多长时间?
24. 若 12 头牛在 4 周内吃掉牧场 3 1/3 英亩的青草,而 21 头牛在 9 周内吃掉同一牧场 10 英亩的青草,问多少牛将在 18 周内吃掉 36 英亩的青草?(提示:青草在不断生长.)
25. 给定直角三角形的周长为 a,面积为 b^2,求其斜边的长度.

麦克劳林《代数学》中的问题

26. 假定伦敦与爱丁堡之间的距离为 360 英里，一位信使以每小时 10 英里的速度从这座苏格兰城市向伦敦急行，与此同时，另一位以每小时 8 英里的速度从这座英国首都向爱丁堡行进. 试问他们将在何处相遇?
27. 从二元一次方程组的克拉默法则推导三元一次方程组的克拉默法则. 给定方程组:

$$ax + by + cz = m$$
$$dx + ey + fz = n$$
$$gx + hy + kz = p$$

从每个方程中解出 x，以 y 和 z 表示之，然后组成后两个变量的两个方程，并解出 z. 最后，通过代入法确定 y 和 x.

28. 一起用餐的一伙人发现账单总额为$175. 其中两个人不得付账. 其余人员发现自己的份额比全部人员一起付账时的份额多出$10. 试问这群人共有多少?

欧拉《代数学引论》中的问题

29. 人们怎样才能调和代数法则 $\sqrt{a}\sqrt{b} = \sqrt{ab}$ 与运算 $\sqrt{-1}\sqrt{-4} = -2 \neq \sqrt{(-1)(-4)}$ 之间的矛盾?你为何认为欧拉在讨论这一情况时犯了错误?
30. 男女共二十人在一个小酒馆一起用餐，账单按每个男人付$8，每个女人付$7 的方法进行分摊，而账单的总数为$145，试求男女各为有多少人?
31. 一名马贩以一定数量的克朗买得一匹马，又以 119 克朗将其卖掉. 如此意味着他所获利润与其购马花费具有相同比例. 买价是多少?
32. 三兄弟以$100 购得一块葡萄园. 老三说，如果老二将其拥有金钱的一半给自己，他就能独自买得葡萄园；老二说，如果老大愿意只将其金钱的三分之一给自己，他就能单独买得葡萄园；最后，老大说，只需要再有老三金钱的四分之一，他也能自己买下葡萄园. 他们每个人各有多少钱?
33. 计算 $1,5,5^2,\cdots$ 以 13 为模的不同剩余 $1,\alpha,\beta,\cdots$. 选择乘幂数列的一个非剩余 x，求其陪集 $x, x\alpha, x\beta,\cdots$. 继续挑选非剩余并求出其陪集，直到关于模 13 的所有 12 个非零剩余类被划分为不相重叠的子集，即 5 的乘幂群的陪集.
34. 求模 13 的二次剩余.
35. 证明：-1 是关于素数 q 的二次剩余，当且仅当 $q \equiv 1 \pmod 4$.

拉格朗日的问题

36. 证明：如果 n 是素数，那么，$x^n - 1 = 0$ 的所有根都可以被表示为其中任何一个根 $\alpha \neq 1$ 的乘幂.
37. 设 x_1, x_2 是二次方程 $x^2 + bx + c = 0$ 的两个根. 由于在根的两种置换下 $t = x_1 + x_2$ 无变化，而 $v = x_1 - x_2$ 取两个不同值，所以 v 必定满足与 t 有关的二次方程. 求出该方程. 类似地，x_1 在对 $x_1 - x_2$ 所实施的相同置换下也无变化，所以，x_1 可以由 $x_1 - x_2$ 的有理式表出. 求出这样一个有理表达式. 利用有理表达式和你找到的方程"解"原二次方程.
38. 求方程 $x^3 - 6x - 9 = 0$ 的三个根 x_1, x_2, x_3. 利用拉格朗日的方法求出 y 所满足的六次方程，其中 $x = y + 2/y$. 求出该方程的所有六个根，并将每一个根明确表示为 $\frac{1}{3}(x' + \omega x'' + \omega^2 x''')$，其中 (x', x'', x''') 是 (x_1, x_2, x_3) 的一个置换，而 ω 是 $x^3 - 1 = 0$ 的一个复根.

克莱罗、欧拉、蒙日的问题

39. 克莱罗揭示了通过积分学计算空间曲线长度的方法，即求 $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}$ 的积分. 利用该结论求柱面 $ax = y^2$ 和 $(9/16)az^2 = y^3$ 的交线介于原点与 (x_0, y_0, z_0) 点之间的长度.

40. 利用克莱罗的方法计算柱面 $x^2 - a^2 = y^2$ 和 $y^2 - a^2 = z^2$ 之交线的次切线和切线.

41. 计算由 $ax = y^2, by = z^2$ 所定义的曲线上的点 P 到 xz- 平面的垂线长度.其中该垂线也垂直于曲线的切线和次切线所确定的平面.

42. 证明平面 $\alpha x + \beta y + \gamma z = a$ 与 xy- 平面之间的夹角 θ 由 $\cos\theta = \gamma / \sqrt{\alpha^2 + \beta^2 + \gamma^2}$ 确定.计算该平面与另两个坐标平面之夹角的余弦值.

43. 证明:如果 $\beta \frac{\partial z}{\partial x} - \alpha \frac{\partial z}{\partial y} = 1$,则平面 $z = \alpha y - \beta x + \gamma$ 与曲面 $z = f(x, y)$ 垂直.(证明该平面包含了曲线的一条法线.)

44. 求平面 $Ax + By + Cz + D = 0$ 的过点 (x_0, y_0, z_0) 的法线.

45. 将蒙日关于曲面 $z = f(x, y)$ 的法线方程形式转化为现代的向量形式.

46. 证明:如果仅有两块区域或没有区域连接奇数座桥,那么,途经一系列与某些区域相连的桥梁的"欧拉线路"(每座桥梁仅途经一次)总是存在的.

47. 在图 14.20 所给情形下,构造欧拉线路.

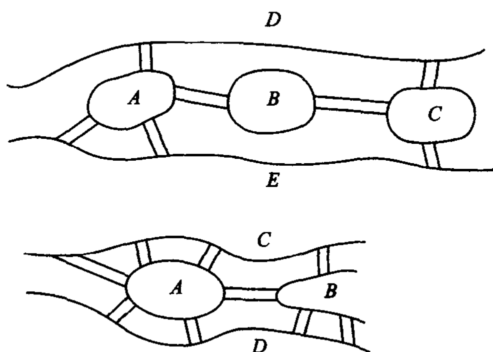


图 14.20　过桥线路问题.

班尼克尔笔记中的问题

48. 班尼克尔很喜欢解决数学难题,并在其笔记本中作了大量记录.其中包括被他改写后的古老的百禽问题:一位绅士给了他的仆人 100 镑,让其购买 100 头家畜,并令其对每头公牛付 5 镑,对每头母牛付 20 先令,对每只羊仅付 1 先令.(20 先令为 1 英镑.)那么,他为自己的主人带回的各类家畜分别有多少头?[47]

49. 将 60 分为四部分,以使第一部分加 4,第二部分减 4,第三部分乘以 4,第四部分除以 4,均等于同一个数.

50. 假设梯子高 60 英尺,被置于街道上以达到街道一侧 37 英尺高的窗户,而在不移动梯子下端的情况下,它可以够着街道另一侧 23 英尺高的窗户.试问街道有多宽?

讨论题

51. 所谓的圣彼得堡悖论(St. Petersburg Paradox)是 18 世纪的数学家们在概率论领域内争论的一个话题.悖论涉及到由两人参与的如下游戏.游戏者 A 抛掷一枚硬币直到其出现反面为止.如果反面在第二次抛掷时出现,B 将赔付 2 卢布,如果到第三次才出现反面,则 B 需赔付 4 卢布,…,如果到第 n 次时才出现反面,则 B 需赔付 2^{n-1} 卢布.问 A 将会愿意付多少钱给 B 以获得抛掷硬币的特权.首先证明 A 的期望值,即游戏的每种可能结果的概率与其收益乘积之和为

$$\sum_{i=0}^{\infty} \frac{1}{2^i} 2^{i-1},$$

而且该和式的值为无穷大.然后试玩 10 次游戏,并计算平均收益.你愿意出多少钱以获得游戏权?为什么数学期望的概念在这个实例中似乎不起作用?

52. 设计一堂统计学课程,来推导贝叶斯定理,并讨论其实用性.

53. 设计一堂代数学课程，利用麦克劳林的方法来讲授克拉默法则的基本原理.
54. 准备一份报告，论述日本数学家关孝和对行列式的发明.
55. 比较麦克劳林和欧拉在其各自的著作中关于正负数乘法运算的论述. 他们各自默认的代数法则是什么？
56. 克莱罗讲授几何学的自然方法是合理有效的教学方法吗？将其与大概源于欧几里得的综合教学方法进行比较研究.
57. 准备一堂三维解析几何课程，利用蒙日的成果推导曲面的切平面方程.

文献和注解

关于概率与统计早期发展史的好书中包括了第 11 章所引用过的 Hacking 与 David 的著作. I. Todhunter, *A History of the Mathematical Theory of Probability from the Time of Pascal to that of Laplace*(New York: Chelsea, 1965)，是一本详细介绍开拓者们的概率论成果的早期著述. Stephen M. Stigler 的 *The History of Statistics*(Cambridge: Harvard University Press, 1986)，对自然包括概率论成就在内的早期统计学的发展史进行了精彩论述. Lorraine Daston 的 *Classical Probability in the Enlightenment*(Princeton: Princeton University Press, 1988)，是一本更加哲学化的论著，它在概率论概念背后的思想方法上集中了较多的注意力. 在新近出版的几部关于代数学史的著作中有个别章节论述了 18 世纪的代数学. 它们是 Lubo s Novy 的 *Origins of Modern Algebra*(Prague: Academia Publishing House, 1973)，Hans Wussing 的 *The Genesis of the Abstract Group Concept*(Cambridge, MIT Press, 1984)，以及 B. L. van der Waerden 的 *A History of Algebra*(New York: Springer-Verlag, 1985). 论述非欧几何学史的权威著作当数 Roberto Bonola 的 *Non-Euclidean Geometry, A Critical and Historical Study of its Development*(New York: Dover, 1955)，由 H. S. Carslaw 翻译. 最近有两部著作出版，均论及萨凯里和朗贝特的有关资料，它们是 Jeremy Gray 的 *Ideas of Space: Euclidean, Non-Euclidean and Relativistic*(Oxford: Clarendon Press, 1989)，以及 Boris A. Rosenfeld 的 *A History of Non-Euclidean Geometry: Evolution of the Concept of a Geometric Space*(New York: Springer-Verlag, 1988)，由 Abe Shenitzer 翻译. Gray 的论文 "Non-Euclidean Geometry—A Re-interpretation" 是对该主题的一个简要研究. 该文见载 *Historia Mathematica* 第 6 卷(1979 年)第 236 – 258 页. 最后，在 Dirk J. Struik 的 "Outline of a History of Differential Geometry" 中可以找到关于微分几何学发展史的综合研究，该文见载于 *Isis* 第 19 卷(1933 年)，92 – 120. 而在 Carl Boyer 的 *History of Analytic Geometry*(New York: Scripta Mathematica, 1956) 详细叙述了解析几何的发展史，其中包括 18 世纪的资料.

1. 摘自 Jakob Bernoulli 的 *Ars Conjectandi* 第四卷，见 James Newman 主编的 *The World of Mathematics*(New York: Simon and Schuster, 1956) 第 3 卷, p. 1452.
2. 同上, p. 1453.
3. 这些公式及相关材料可以在伯努利 *Ars Conjectandi* 的第二卷中找到，见 D. J. Struik, *A Source Book in Mathematics, 1200—1800*(Cambridge: Harvard University Press, 1969). 对 *Ars Conjectandi* 全文的一个简要考察来自 Ian Hacking 的 "Jacques Bernoulli's Art of Conjecturing"，文载 *British Journal for the History of Science* 22(1971), 209 – 229.
4. 参见 Stigler, History of Statistics, pp. 66 – 70. 更详细的资料也可参见 O. B. Sheynin 的 "On the Early History of the Law of Large Numbers"，文载 *Biometrika* 55(1968), 459 – 467，并被收入 E. S. Pearson 和 M. G. Kendall 的 *Studies in the History of Statistics and Probability*(London: Griffin, 1970), 231 – 240；以及 Karl Pearson 的论文 "James Bernoulli's Theorem"，文载 *Biometrika* 17(1925), 201 – 210.
5. 有关 18 世纪保险、养老金和彩票的更详细资料，请参阅 Daston, *Classical Probability* 第 3 章.
6. Abraham De Moivre, *The Doctrine of Chances*, 3rd ed. (New York: Chelsea, 1967), p. 36. 全书的再版因其问题丰富、事例详实而很值得一读. 关于棣莫弗的更多信息可以查 H. M. Walker 的 "Abraham De Moivre"，文载 *Scripta Mathematica* 2(1934), 316 – 333，以及 Ivo Schneider 的 "Der Mathematiker Abraham De Moivre(1667—1754)," 文载 *Archive for History of Exact Sciences* 5(1968), 177 – 317.
7. 同上, p. 37.
8. 同上, p. 242.

9.被引用于 Stigler, *History of Statistics*, p.76.

10.De Moivre, *Doctrine of Chances*, p.246.

11.同上,p.248.

12.Thomas Bayes,"An Essay towards solving a Problem in the Doctrine of Chances",连同 G.A.Barnard 所写的一个作者简介,可参见 E.S.Pearson 和 M.G.Kendall, *Hisory of Statistics ard Probability*, 131 – 154, p.136.

13.同上,p.139.

14.同上,p.140.

15.同上,p.143.

16.同上.该段引文曾引发出对贝叶斯定理适用情形的众多争论,针对该问题最近有两篇分析,其一是 Stigler, *History of Statistics*, pp.122 – 131,其一是 Donald A.Gillies 的"Was Bayes a Bayesian?",见 *Historia Mathematica* 14(1987),325 – 346.

17.Isaac Newton, *The Mathematical Papers* 5, p.65.

18.同上,p.73.

19.同上,p.137.

20.同上,p.141.

21.Colin Maclaurin, *A Treatise of Algebra in Three Parts*, 2nd ed.(London:Millar and Nourse,1756),p.1.

22.同上,p.132.

23.Leonhard Euler, *Elements of Algebra*,由 John Hewlett 译自 *Vollständige Anleitung zur Algebra*(New York:Springer-Verlag,1984),p.2.该 Springer 版本翻印自 1840 年的英文译本,附有 C.Truesdell 所写的序言.其中许多问题以及独创性的解法很值得仔细阅读.

24.同上,p.186.

25.同上,p.7.

26.同上,p.43.

27.同上,p.91.

28.同上,91 – 92.

29.同上,p.413.

30.引自 Jean-Luc Dorier 的"A General Outline of the Genesis of Vector Space Theory",文载 *Historia Mathematica* 22(1995),227 – 261,p.230.欧拉的原文名为"Sur une contradiction apparente dans la doctrine des lignes courbes",见 *Mémoires de l'Académie des Sciences de Berlin*,4(1750),219 – 223.另见 *Opera Omnia* (3)26,33 – 45.

31.Alexis-Claude Clairaut, *Éléments de géométrie*(Paris,1741),绪言.

32.Girolamo Saccheri, *Euclides Vindicatus*,由 G.B.Halstead 翻译(New York:Chelsea,1986),p.9.该版本是对 1920 年的第一个英文译本的再次印刷,另加入了 Paul Stäckel 和 Friedrich Engel 的注.要了解对萨凯里思想的简要论述,请参阅卡 Louis Kattsoff 的"The Saccheri Quadrilateral",见 *Mathematics Teacher* 55(1962),630 – 636.

33.同上,p.117.

34.同上,p.169.

35.同上,p.173.

36.John Fauvel 和 Jeremy Gray 主编, *The History of Mathematics: A Reader*(London:Macmillan,1987 年),517 – 520,p.517.该原典读本选入了朗贝特《平行理论》的一段英文节选.要更多地了解朗贝特,可参阅 J.J.Gray 和 L.Trilling 的论文"Johann Heinrich Lambert, Mathematician and Scientist",见 *Historia Mathematica* 5(1978),13 – 41.

37.同上.

38.同上,p.520.

39.Alexis-Claude Clairaut, *Recherches sur les courbes á double courbure*(Paris:Quillau,1731),p.39.

40.Leonhard Euler,"Methodus facilis omnia symptomata linearum curvarum non in eodem plano sitarum investigandi," *Acta Acad. Sci. Petrop*. 1(1782),19 – 57 = *Opera*(1)28,348 – 381.

41. Leonhard Euler: "Recherches sur la courbure des surfaces", *Mem. de l'Academie des Sciences de Berlin* 16(1760), 119 - 143 = *Opera*(1)28, 1 - 22.

42. 引自 Boyer 的 *Analytic Geometry*, 205 - 206.

43. Norman L. Biggs, E. Keith Lloyd, Robin J. Wilson, *Graph Theory: 1736—1936*(Oxford: Clarendon Press, 1986), p.6. 该卷收录了欧拉关于哥尼斯堡七桥问题的论文的全译本.

44. Janis Langins, *La République avait besoin de savants*(Paris: Belin, 1987), p.123. 该书详细研究了综合工科学校创办第一年的情况,并提供了许多相关文件的副本.

45. 引自 Dirk J. Struik 的"Mathematics in Colonial America",见 Dalton Tarwater 主编的 *The Bicentennial Tribute to American Mathematics: 1776—1976*(Washington: Mathematical Association of America, 1976), 1 - 7, p.3.

46. 引自 David Eugene Smith 和 Jekuthiel Ginsburg 的 *A History of Mathematics in America before 1900*(Chicago: Mathematical Association of America, 1934), p.62.

47. 参阅 Silvio A. Bedini, *The Life of Benjamin Banneker*(New York: Scribner's, 1972),以获取有关班尼克尔生平及其工作的详细资料,此外还有相关文件的副本,包括此处收录的问题.

18 世纪概率论、代数学、几何学概览

1642—1708	关孝和(Seki Takakazu)	行列式
1642—1727	牛顿(Isaac Newton)	通用算术
1654—1705	伯努利(Jakob Bernoulli)	组合学与概率论
1667—1733	萨凯里(Girolamo Saccheri)	非欧几何
1667—1754	棣莫弗(Abraham De Moivre)	概率论、正态曲线
1698—1746	麦克劳林(Colin Maclaurin)	代数课本、克拉默法则
1702—1745	格林伍德(Isaac Greenwood)	哈佛大学教授
1702—1761	贝叶斯(Thomas Bayes)	概率论、贝叶斯定理
1707—1783	欧拉(Leohard Euler)	代数、数论、几何、拓扑
1713—1765	克莱罗(Alexis Clairaut)	几何、空间曲线
1714—1779	温思罗普(John Winthrop)	哈佛大学教授
1728—1777	朗贝特(Johann Lambert)	非欧几何
1731—1806	班内克尔(Benjamin Banneker)	测量学、年历
1732—1796	黎滕霍斯(David Rittenhouse)	天文学、测量学、钟表制造
1736—1813	拉格朗日(Joseph Lagrange)	方程理论
1746—1818	蒙日(Gaspard Monge)	分析与微分几何
1749—1827	拉普拉斯(Pierre Simon de Laplace)	统计推断
1752—1833	勒让德(Adrien Legendre)	数论
1765—1843	拉克鲁瓦(Sylvestre Lacroix)	教科书

第15章　19世纪的代数

> 实整数可以分解成素数的乘积,这个性质对任意一个给定的整数都成立.但是,令人遗憾的是,实整数的这个性质对于复整数(任意割圆域)并不成立.否则的话,当前我们仍面临许多困难的理论体系就可以很容易解决并得出结论了.因此,看起来我们一直考虑的复整数并不十分完美.这就提出了一个问题:是不是可以找到与这种实整数相似的另一类复整数,它保持了实整数的基本性质.
>
> (库默尔,摘自 *De numeris complexis, qui radicibus unitates et numeris integris realibus constant*,1847)[1]

在暮秋的一个下午,威廉·罗万·哈密顿发现了四元数,正如他后来给他儿子的信中所写到的:"在16号那天(1843年10月)—— 碰巧是个星期一,爱尔兰皇家学院召开会议的那一天 —— 我走在去参加并主持这个会的路上,你母亲陪着我沿着皇家运河走 …… 尽管她一路上不时地与我交谈,一种潜意识还是出现在我的脑海里,并最终有了结果.不用说,关于这个结果,我马上意识到了它的重要性:电路接通了,闪出火花 …… 我无法抑制住激动的心情 —— 也许是不理智的:用刀子把它刻在了布朗汉姆桥上.这个用符号 i,j,k 表示的基本公式就是:

$$i^2 = j^2 = k^2 = ijk = -1,$$

它包含了上述问题的解."[2]

代数学在19世纪主要是研究解方程.到了20世纪,它也包含了各种数学结构的研究,即那些具有精确定义的运算规则的元素的集合,这些元素满足某些特定的公理.代数学在这方面的发展就是我们本章所要讨论的主题.

19世纪的研究是以高斯《算术研究》的出现为开端的.在这本书里,这位"数学王子"讨论了数论的基础,不仅证明了二次互反原理,并且引入了各种新的观念,提供了最早的群和矩阵的例子,高斯的关于高阶互反律的研究很快引出了他对所谓高斯整数的研究,这种数的形式为:$a + bi$,其中 a, b 为普通的整数.把这类整数的性质向其它数域推广的尝试使得库默尔意识到这些重要性质中的某些性质,比如说,惟一因子分解性质是不成立的,为了得到这个性质,使用了术语"素"的一个

引申义.在 1946 年创造了他自己称之为"理想复数"的数.库默尔在理想复数方面的研究导致了戴德金 19 世纪 70 年代定义了代数整数环上的"理想".这些理想能够惟一分解为一些素元之积.

高斯在他的《算术研究》中关于割圆方程的解的论述和柯西在 1815 年关于置换的详细论述对高于四次的代数方程解的问题有新的帮助.最终是阿贝尔证明五次或更高次的代数方程不可能有一般的根式解(在 1827 年).之后不久,伽罗瓦概述了代数方程和根的置换之间的关系.该关系的完全发展把方程的可解性问题转化为因子群和方程群子群的研究,伽罗瓦的著作直到 1846 年才出版,被完全理解是更晚些的时候了.在 1854 年,凯莱给出了抽象群的最早定义,但是他的论述在当时那个时代同样有点超前,直到 19 世纪 70 年代他本人和更晚一些时候迪克和韦伯才对此进行了更深入的研究.同时,为了解代数方程而对数进行的研究导致了克罗内克和戴德金对数域的定义和之后不久韦伯对域的抽象定义.

19 世纪前 30 年的英国数学家们,包括皮科克和德摩根致力于代数学最基本的思想的公理化,并确定整数的性质究竟有多少可以推广到其它形式的量上去.这些研究最终使得哈密顿在 1843 年发现了四元数,其中部分是由于他尝试在三维空间中定义有物理意义的代数.然而,四元数是四维的对象,因此,物理学家在任何代数运算中只能应用其中的三维部分.在持续到接近世纪末的一场争论后,由奥列维·亥维塞和约西阿·魏拉尔德·吉布斯发展的四元数代数占了上风,并使之成为物理学家的语言.同时,哈密顿为了物理应用而倡导的可以自由确定代数学中的运算法则的观念,被布尔在逻辑学的研究中所引用.他的研究对一个世纪之后计算机的设计起了很重要的作用.

另一部分在 19 世纪中期得到发展的方向是我们今天称之为近世代数学的矩阵论的研究.行列式早在 17 世纪就被使用了,但是,直到 1850 年詹姆斯·约瑟夫·西尔维斯特才用矩阵来表示矩形的数字阵列.之后不久,凯莱发展了矩阵代数.柯西在本世纪初关于二次型的工作中就开始了特征值的研究,而后约当,弗罗贝尼乌斯以及其他许多人发展了柯西的思想,尤其是弗罗贝尼乌斯把矩阵理论整理成了本质上今天的形式.

15.1　数　论

勒让德于 1798 年发表了他在数论方面的工作,同时,在德国北部不伦瑞克城的一位年轻人也即将完成他的数论工作.这部著作所产生的影响比勒让德的著作还要深远的多.高斯(1777—1855)在他 1801 年出版的《算术研究》的序言中提到,在研究同时代的其他人的著作之前,《算术研究》的大部分就已经写成了.尽管欧拉、拉格朗日或勒让德也已经知道高斯发现的部分理论,但其著作中包含了许多数论方面的新发现.

15.1.1　高斯与同余

高斯的第一章首先叙述了同余的现代定义和符号:如果 a 整除 b 与 c 的差,则说整数 b 和 c 相对于**模** a 是**同余的**.高斯把它表示为:

$$b \equiv c \pmod a.$$

"由于注意到等于与同余之间的类似,"[3] 他采用了符号"$\equiv$",并称之为 b 与 c 互为**余数**.之后他讨论了同余的基本性质.比如,高斯证明了如果 a 与 m 的最大公约数是 1 的话,由欧几里得算法知,线性同余式 $ax + b \equiv c \pmod m$ 是可解的.他还给出了如何解中国余数问题,以及欧拉函数 $\phi(n)$ 的计算,$\phi(n)$ 是指小于 n 并与 n 互素的整数个数.在第 3 章,就像在他之前的欧拉一样他考虑了幂的余

数，提到如果 p 是一个素数，a 是小于 p 的任何一个数，那么使得 $a^m \equiv 1 \pmod p$ 成立的最小的 m 是 $p-1$ 的一个因数.继而，高斯指出"事实上总有一些数，前 $(p-1)$ 次幂是不同余 1 的，"[4] 具有这种性质的数 a 叫做模 p 的一个**原根**，如果是 a 一个原根那么 a 的方幂 $a^2, a^3, \cdots, a^{p-1}$ 模 p 的余数互不相同，取遍了 $1,2,\cdots,p-1$.特别地，$a^{\frac{p-1}{2}} \equiv -1 \pmod p$.这是高斯在

威尔森定理：$(p-1)! \equiv -1 \pmod p$

的证明中用到的一个重要性质.令 a 是模 p 的一个原根，那么，

$$(p-1)! = a^1 a^2 \cdots a^{p-1} = a^{1+2+\cdots+p-1} = a^{p(p-1)/2}.$$

因为 $p(p-1)/2 \equiv (p-1)/2 (\bmod\ p-1)$，由此可得：

$$(p-1)! \equiv a^{(p-1)/2} \equiv -1 \pmod p.$$

这个理论最早是被 J.威尔森(1741—1793)提出的.后来由拉格朗日在 1773 年首先证明.

《研究》第 4 章的主题是二次互反原理.由欧拉提出但未被证明.这个原理给出了两个奇素数二次互余的条件.高斯把这个定律看的非常重要以至于给出了六种不同的证明.第一种证明是经过艰苦的努力之后在 1796 年春天发现的.在《研究》里，高斯在他的第二种证明给出之前列出了从计算中获得的大量的例子和特殊情况.在欧拉基础之上，他先得出 -1 是形为 $4n+1$ 的素数的二次剩余，但不是形为 $4n+3$ 的素数的二次剩余.下一步，他讨论了 2 和 -2.并得出结论：对于形如 $8n+1$ 的素数来说，2 和 -2 都是它的二次剩余.对于形如 $8n+3$ 的素数来说 2 不是，-2 是.对于形如 $8n+5$ 的素数来说 2 和 -2 都不是它的二次剩余.对于形如 $8n+7$ 的素数来说 2 是 -2 不是.这些性质的证明并不难.比如在证明第一个结果时，高斯选了形为 $8n+1$ 的素数的一个原根 a.注意到 $a^{4n} \equiv -1 (\bmod\ 8n+1)$ 可以重写为下面两种形式中的任何一种：

$$(a^{2n}+1)^2 \equiv 2a^{2n} (\bmod\ 8n+1) \quad \text{或} \quad (a^{2n}-1)^2 \equiv -2a^{2n} (\bmod\ 8n+1),$$

因此，$2a^{2n}$ 和 $-2a^{2n}$ 平方同余 $(8n+1)$，但是由于 a^{2n} 是一个平方式，因此 2 和 -2 也平方同余 $(8n+1)$.高斯继续刻画 ± 3，± 5，± 7 是 2 次剩余的素数的特征，这样，他就可以叙述一般的结论：

二次互反律：如果 p 是形如 $4n+1$ 的一个素数，那么 p 是不是一个素数的二次剩余在于该素数是不是 p 的余数，对于形如 $4n+3$ 的素数 p，$-p$ 也有同样的性质.

上述定理的证明很长，在此不予讨论，但是应特别指出高斯没有注意到勒让德关于二次余数的有启发性的符号，这个符号可以很明了地表示出该结果.勒让德定义了符号 $\left(\frac{p}{q}\right)$，当 p 是 mod q 的二次剩余时，$\left(\frac{p}{q}\right) = 1$，否则，为 -1.式中的 q 是一个奇素数，于是这个定理就可以用以下工整的形式表示：

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

-1 和 ± 2 模素数 p 的二次余数的性质也可以用类似的公式表示，下面给出余数的乘法规则：$\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = \left(\frac{ab}{p}\right)$ 式中 a，b 与 p 互素，同样，当两个数有公因子时，也给出了用以确定余数的情况的规则，后来高斯对任意两个正数 P、Q.无论 Q 是不是 P 的余数都给出了定义，比如说，看 453 是不是模 $1236(=4\times3\times103)$ 的二次剩余.首先，注意到 453 是 4 和 3 的二次剩余，由中国剩余定理这个问题就转化成了确定 $\left(\frac{453}{103}\right)$，应用勒让德的符号，我们有

$$\left(\frac{453}{103}\right) = \left(\frac{41}{103}\right) = \left(\frac{103}{41}\right)(-1)^{\frac{41-2}{2}\frac{103-1}{2}} = \left(\frac{103}{41}\right) = \left(\frac{-20}{41}\right)$$

$$= \left(-\frac{1}{41}\right)\left(\frac{2^2}{41}\right)\left(\frac{5}{41}\right).$$

因为右端三个因子的每一个都等于 1,所以 453 是 103 和 1236 的二次剩余,事实上,高斯得出 $453 \equiv 297^2 \pmod{1236}$.

人物小传	卡尔·弗里德里希·高斯(1777—1855)(Carl Friedrich Gauss)
	高斯(图 1.5)出生在一个刚从农村搬到城市的一个家庭里,就如当时的许多家庭一样,希望以此来改善作为贫苦的农业工人的命运.在不伦瑞克(Brunswick)生活的好处就是年幼的高斯可以去上学.有许多关于他早慧的故事,其中一个发生在他 9 岁时的数学课上,当时新学年刚刚开始,彼特奈尔老师为了使他的 100 多位学生集中精力上课,要求他们把前 100 个数加起来.他刚布置完任务,高斯就把答案 5050 写在了他的石板上交给了老师.他注意到问题的答案就是 50 个 101 相加,即 101 个不同的数对 1 与 100,2 与 99,3 与 98…… 的和,他心算算出了乘法.这位年幼的学生给他留下了深刻的印象,他为高斯安排了特殊的教材,并让他的助手马丁·巴泰尔斯(1769—1836)(他后来成为俄国的一位数学教授)专门指导高斯.并让他直接进入到中学学习,在中学里,他很好地掌握了那些经典课程. 在 1791 年,不伦瑞克的公爵为高斯提供了奖学金,使他可以到卡罗林纽学院学习.这是一所为了培训政府官僚和军官而由镇政府出资而建的研究院.这个研究院以新科学为主导.后来高斯又被录取到相邻的汉诺威的哥廷根大学继续深造,这所大学在当时的科学界很有声望.最后他又返回到不伦瑞克继续他的研究.同时在当地的海尔姆斯威特大学获得博士学位.在 1801 年,高斯除了发表他在数论方面的发现外——为了报答公爵,他将著作献给了公爵,同时提出了一种新的计算轨道的方法,用这种方法又发现了几颗小行星.公爵的资助一直持续到 1806 年,公爵在这一年的对法战争中牺牲,公国也被法国占领.法国将军收到了对高斯的生命财产给予照应的明确命令,这真是科学界的一大幸事,因而他得以在不伦瑞克一直呆着,直到第二年他以天文学教授和天文台的主任身份在哥廷根大学得到一个职位.此后,他一直孜孜不倦地做着纯粹的研究:应用数学,天文学,大地测量学…… 高斯并不十分喜欢给学生讲课,因为许多学生不热爱数学,甚至基础还特别差.但是他非常喜欢私下和那些接近他并真正对数学感兴趣的学生一起交流.与他的前辈以及同时代的法国人相比,高斯出版的东西很少.他的作品集只有 12 卷.然而,他在各个领域的数学论文是如此的深刻,以至于一直影响到这门学科今天的发展.[5]

在后来的几十年里,高斯想要把二次互反律推广到三次和四次互反律,也就是确定什么时候一个数模其他数 3 次和 4 次同余的定律.早在 1805 年他就意识到"二次同余的一般理论是算术领域的扩充的源泉."[6] 这就导致了他对形如 $a+bi$ 的复数的思考,其中 a,b 均为普通整数.这种复数被称为高斯整数.他对这方面的研究含在他 1832 年发表的论文中.在这篇文章中,他对得到的高斯整数和通常整数作了类比,注意到高斯数有单位元素(可逆元)$1,-1,i,-i$ 后,他定义了整数的**模**为 a^2+b^2,即它与它的共轭复数 $a-bi$ 的乘积.他把那些不能表达成另外两个非单位元素整数的乘积的整数叫做**素数**.它的下一个任务是确定哪些高斯数是素数,因为一个实的奇素数可以写成形式 $p=a^2+b^2$ 当且仅当它是形如 $4n+1$ 的数.因此,这种高斯数是复合数,写成 $p=(a+bi)(a-bi)$.相反地,形如 $4n+3$ 的素数看作高斯



图 15.1　德国邮票上的高斯.

数时仍是素数.因为 $2=(1+i)(1-i)$,2是复合数,为了确定其余的哪些是素数,高斯用了模的概念,并指出 $a+bi$ 是素数当且仅当它的模是一个实素数.即只能是2或形如 $4n+1$ 的数,换句话说就是2和形如 $4n+1$ 的素数在高斯整数环内可以拆成两个素数的乘积,形如 $4n+3$ 的素数在这里仍然是素数.

利用高斯整数环中的素数,高斯证明了算术基本定理的类比定理,其证明用了与欧几里得的《原本》第七册中相似的定理.首先他很容易地证明了任何整数都可以分解成一些素数的乘积,为了与一般的整数的性质吻合,紧接着他证明了这种分解至少在相差单位因子的意义上是惟一的.先用它对素数的描述证明:如果一个高斯素数可以分解成一些高斯素数 $q,r,s\cdots$ 的乘积的话,那么一定等于其中的某一个或者某一个与单位元的乘积.在确定了高斯整数的惟一分解性质之后,高斯继续考虑了高斯整数意义下的同余以及二次互反律等性质.更进一步,他发现了三次互反律应包含形如 $a+b\omega$ 的复数有关,在这里 $\omega^3=1$,其中 a,b 是一般的整数.这就需要探讨这种形式的素数以及分解问题.这项工作高斯没能完成.

15.1.2 费马大定理和惟一分解

在各种环中的完全分解的整个问题被证明不仅与互反性有关,而且与费马大定理的攻克有关.费马大定理是说 $x^n+y^n=z^n$ 在 $n>2$ 没有非平凡的整数解.这个定理的 $n=3$ 的情况是由欧拉首先在1753年用费马的无穷递降证明的,发表在他1770年出版的《代数学》中.

下一位在整个定理中做出进展的数学家是杰曼(1776—1831).在19世纪20年代初,她证明了当 xyz 不能被100以内的任何奇素数整除时无解,不幸的是它的方法不适用于在 x,y,z 中有一个可被 n 整除的情况下的费马大定理的情形,在1825年勒让德完整地证明了 $n=5$ 时的情形,七年之后,狄利克雷(1805—1859)证明了 $n=14$ 时的定理的情形.又过了七年之后,G.拉梅(1795—1870)证明了 $n=7$ 的情形,后面几种情形的证明都非常长,使用了非常困难而巧妙的处理,很难推广,看来如果要证明这个定理,必须要找到一个全新的方法.

人物小传	苏菲·杰曼(1776—1831)(Sophie Germain)
	尽管由于法国革命造成的混乱局面和父母的反对使得杰曼不得不秘密地学习,杰曼仍然通过自学微积分而精通了数学.1794年,当综合工科学校设立的时候,她非常想去继续学习.但是,学校不接收女生.尽管如此她仍然孜孜不倦的收集和学习各个数学方面的讲稿和笔记.甚至把她本人的文章递交给拉格朗日看.在高斯的《研究》出版之后,她很快就精通了,并且开始用假名布兰克与高斯通起了信.事实上,是她建议1807年带领军队占领不伦瑞克的法国将军保证高斯的安全的.当然,当时高斯甚至不知道杰曼这个名字,但是这点误解在他们互通信息之后很快就明白了.也许对于在德国长大的数学家高斯来说,当他发现他的通信者以及保护者是一位女性时是非常惊奇的.如他所写:"由于我们的传统和偏见,一位女性使自己熟悉并习惯于那些困难的研究时一定遇到了比男性多得多的困难,尽管如此,她成功地超越了那些晦涩的困扰并能洞察出其中的一些东西,毫无疑问,她一定有高贵的勇气,非凡的才能和过人的智慧."[7]

在1847年3月1日的巴黎学术会议上,拉梅(Gabriel Lamé,1795—1859)宣称他对证明费马大定理有了新的进展.[8] 拉梅称他已解决了这个冗长而复杂的问题,并给出了他的一个简证,它的基本思想始于表达式 x^n+y^n 在复数域上的分解:

$$x^n + y^n = (x + y)(x + \alpha y)(x + \alpha^2 y)\cdots(x + \alpha^{n-1}y),$$

其中 α 是$x^n - 1 = 0$的一个本原根.拉梅进一步想要说明如果 x 和y 使表达式中的因子都是互素的话,那么每一个因子都必然是 n 次幂的,对于表达式也是同样的,然后再利用费马的无穷递降的技巧,在小一些的数字中找到一组解,另一方面,如果这些因子不是互素的,他希望证明它们有一个公因子,用这个公因子去除式子两边,就把问题化简为第一种情况了,当拉梅讲完他的证明思路之后,J.刘维尔(1809—1882) 指出了的证明中的一些非常值得怀疑的地方.本质上,他指出"由于每个因子是互素的,并且它们的乘积是 n 次幂的,因而就有某个因子是 n 次幂的."拉梅的这一推导思想是建立在任何整数都能被惟一分解成一些素数之积的基础上的,但是,这个结论对形如 $x + \alpha^j y$ 的复数来说并不是明显成立的.

在紧接着的几周时间里,拉梅想尽量解决刘维尔提出的问题,但是没有成功,直到 5 月 24 日 ,刘维尔仔细地看了库默尔(1810—1893) 发表在学报上的信后,才有效地结束了这场争论.即库默尔不仅指出在这个问题中的某些地方惟一分解定理是不成立的,并且在 3 年之前他就在一本有些深奥难懂的出版物上发表了一篇文章,在这篇文章中他就阐明了这种错误.库默尔在 1844 年发表的这篇文章是关于高阶互反律问题的,但也包含了与对单位根有关的复数的一般理论的研究.

15.1.3　库默尔和理想数

库默尔所研究的这种与费马大定理和一般互反律有重要联系的复数被称为**割圆整数**,它们是形为:

$$f(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \cdots + a_{n-1}\alpha^{n-1}$$

的复数,式中 α 是$x^n - 1 = 0$的不等于 1 的根,每个 a_i 都是一个通常的整数.特殊地,我们假定这里的 n 是一个素数.用方程 $x^n - 1 = 0$ 的其他不等于 1 的根 α^i 去替换上式中的所形成的数,记作 $f(\alpha^2)$、$f(\alpha^3)\cdots f(\alpha^{n-1})$,称为给定数的**共轭**,所有共轭的乘积称作 $f(\alpha)$ 的**模**,记作 $Nf(\alpha)$,是一个通常的整数,这个模满足关系式,$N[f(\alpha)g(\alpha)] = Nf(\alpha)Ng(\alpha)$,模的这种乘法性质成为库默尔处理割圆整数分解问题的一个基本工具,因为整数 $h(\alpha)$ 的一个分解蕴含通常整数 $Nh(\alpha)$ 的一个因子分解.

为了论述库默尔的证明,我们引入两个定义:如果在某个特定的环中一个复整数不能被分解成该环中其他两个非单位整数的乘积的话,就称为**不可约**的.如果一个数整除一个乘积就整除它的一个因子的话就称之为"**素**"(注意这个定义与高斯的定义不同),很容易证明素元是不可约的,但是库默尔发现,在由一个 23 阶的单位根 α 生成的整环 Γ 中,存在非素元的不可约数.这样,他证明了在环 Γ 上惟一分解不成立.

首先,库默尔证明了 Γ 中的任何整数的模都具有形式$\dfrac{x^2 + 23y^2}{4}$,由此可得,素数 47 和 139 不是任何整数的模,因为无论是 $4 \times 47 = 188$ 还是 $4 \times 139 = 556$ 都不能写成一个数的平方加上另一个数的平方的 23 倍的形式,另一方面,47×139 是 $\beta = 1 - \alpha + \alpha^{-2}$ 的模;由之可得 β 整除 47×139.如果 β 是一个素数,那么它必须整除 47 或 139.但这是不可能的.因为 $N(\beta)$ 即不能整除 $N(47) = 47^{22}$,也不整除 $N(139) = 139^{22}$.如果 β 可被因子分解,那么它的某个因子的模就一定是 47,与 47 不是模相矛盾.因此,β 只能是一个非素元的不可约数.47×139 可显式表示成两个不同的不可约数的乘积.首先,$47 \times 139 = N(\beta)$ 给出了模为 47×139 的 22 个不可约因子的因式分解.第二,如果

$$h(\alpha) = \alpha^{10} + \alpha^{-10} + \alpha^8 + \alpha^{-8} + \alpha^7 + \alpha^{-7} \quad 且 \quad g(\alpha) = \alpha^{10} + \alpha^{-10} + \alpha^8 + \alpha^{-8} + \alpha^4 + \alpha^{-4},$$

那么 $N(h(\alpha)) = 47^2$, $N(g(\alpha)) = 139^2$. 因此 $h(\alpha)$、$g(\alpha)$ 和它们所有的共轭都是不可约的,继而,设 $f(\alpha) = h(\alpha)g(\alpha)$ 可以证明:

$$47 \times 139 = f(\alpha)f(\alpha^4)f(\alpha^{-7})f(\alpha^{-5})f(\alpha^3)f(\alpha^{-11})f(\alpha^2)f(\alpha^8)f(\alpha^9)f(\alpha^{-10})f(\alpha^6),$$

式中的因子是由 α 变成 α^4 而生成的共轭元素, 47×139 的这个新的 22 个不可约因子的分解很明显与原先的那个不同,其中一半因子的模是 47^2 另一半因子的模是 139^2. 非常有趣的是,刘维尔在 1847 年 6 月见到库默尔的证明之前, 1847 年 5 月在他的笔记中就已写下了一个非常简单的非惟一分解的例子,该例子是在由 $x^2 = -17$ 的一个整域中,而不是 $x^n = 1$:

$$169 = 13 \times 13 = \left(4 + 3\sqrt{-17}\right) \times \left(4 - 3\sqrt{-17}\right).$$

在发现了非惟一因子分解之后,在接下来的一年里构造了本章开头部分引用的问题的一个解,他构造了一种新复数:"理想复数".它能够被惟一分解成理想素元的乘积.例如,考察由 23 阶单位根所生成的整环,在这里, 47 和 139 都没有素因子.因为 $N(1 - \alpha + \alpha^{-2}) = 47 \times 139$,但是在"理想素元"的概念里,在式子左边的22个不可约因子中任何一个都应被式子右边的两个"理想素元"整除,一个是47,另一个是139.为了描述这样的素因子,库默尔定义了**整除**:假设 P 是整除 $\beta = 1 - \alpha + \alpha^{-2}$ 的一个 47 的素因子, ψ 是 β 的 21 个共轭的乘积,那么, ψ 将被 P 之外的 47 的所有理想素因子整除,因此, $\gamma\psi$ 被 47 整除当且仅当 γ 被 P 整除,于是当 $\gamma\psi$ 被 47 **整除**时, γ 就被 P 整除(γ 是 23 阶单位根生成的环中的一个整数),类似地,当 $\gamma\psi^m$ 被 47^m 整除时, γ 被 P^m 整除,把这种思想推广到割圆整数的任何环上.又因为任意一个理想数被形式地定义成理想素元的乘积,它成功地把素因子的惟一因子分解建立(实际上是定义)在了理想数上.通过对这种理想数的更深入细致的研究,库默尔能对素数 n 添加某些条件,使得费马大定理在这些 n 成立,并且证明了 100 以内的除了 37、59、67 以外的所有素数的费马大定理.[9]

库默尔在他早期的关于理想复数的论文中写到,他的目的是把他的成果推广到割圆整数之外的整环,即由 $x^2 - D = 0$ 的一个根所生成的整环上去,式中 D 是一个整数,他没有发表任何这种推广.部分是由于他没能发现合适的把整数观念推广到那些整环上的方法.较明显的推广看上去好像是把整环中的数写成形如 $a + b\sqrt{D}$ 的复数形式,这里 a、b 是一般整数.但是,这种定义引出一些问题,比如,在形如 $a + b\sqrt{-3}$ 的数形成的整环中,令 $\beta = 1 + \sqrt{-3}$,那么 $\beta^3 = -8$.既然 2 在该环中不能整除 β,就必然存在2的一个理想素因子 P.他除2得的重数比除 β 的重数大,为简单起见,我们把它记为 $\mu_P(2) > \mu_P(\beta)$,另一方面,对任意的 k, 2^k 整除 $8\beta^k$,因此对任意的 k,有 $k\mu_P(2) \leqslant \mu_P(8) + k\mu_P(\beta)$,即 $k(\mu_P(2) - \mu_P(\beta)) \leqslant \mu_P(8)$ 成立,这就意味 $\mu_P(2) \leqslant \mu_P(\beta)$ 矛盾.

人物小传	恩斯特·库默尔(1810—1893)(Ernst Kummer)
	库默尔,生于德国索奥市(现在波兰的扎里)位于柏林和弗罗茨瓦夫(过去的布雷斯劳)正中间,于 1828 年进入哈雷大学学习神学,但他很快把专业方向转到了数学,于 1831 年获得博士学位以后,在利格奈特(现在的利格奈卡)的一个高级中学里教了十年的数学与物理. 1842 年,他被聘请到布雷斯劳大学,正是在这里,库默尔把他的一生奉献给了数论研究. 1855 年,在狄利克雷离开柏林去哥廷根接替高斯的位置的时候,库默尔被任命补了他在柏林的空缺,在这里,他与魏尔斯特拉斯一起设立了德国第一个纯数学讨论班,这个讨论班很快吸引了世界各地的数学家.并使柏林在 19 世纪末 20 世纪初成为世界几大数学中心之一.

15.1.4 戴德金和理想

整数的定义问题是由戴德金(1831—1916)在 1871 年解决的,此外,由于不满"库默尔没有定义

理想数本身,只用他们定义了整除性,"[10] 戴德金创立了一种新的把惟一因子分解恢复到他刚定义的整环上去的新概念.戴德金首先定义了**代数数**为任何满足有理数上的代数方程的复数,此方程形如:

$$\theta^n + a_1\theta^{n-1} + a_2\theta^{n-2} + \cdots + a_{n-1}\theta + a_n = 0,$$

式中的 a_i 为有理数.然后,他定义当满足上面的方程中的系数全为通常整数时,其代数数为代数整数.比如 $\theta = \frac{1}{2} + \frac{1}{2}\sqrt{-3}$ 是一个代数整数,因为它满足方程 $\theta^2 - \theta + 1 = 0$,尽管它不是"显然"形如 $a + b\sqrt{-3}$(a、b 为整数)的数,然后戴德金证明了代数整数的和、差、积仍是代数整数.他以标准的形式的定义了整除性:"一个代数整数 α 能被一个代数整数 β **整除**,如果 $\alpha = \beta\gamma$ 对于某一个代数整数 γ 成立.但是为了发展一般的整除性,他需要把他的研究限制在代数整数环的一部分上.因此任给一个满足 n 次不可约方程的代数整数 θ,他定义对应于 θ 的代数整数系 Γ_θ 为形如 $x_0 + x_1\theta + x_2\theta^2 + \cdots + x_{n-1}\theta^{n-1}$ 的代数整数的集合,上式中的 x_i 为有理数,像库默尔一样戴德金现在可以在整环 Γ_θ 中定义什么整数是素元或者不可约的了.

戴德金证明了高斯整数 Γ_i 是一个代数整数系,高斯证明了惟一素元分解在这个系中成立.但是,根据狄利克雷研究,他提供了一个不同于高斯的证明,即他首先证明了任意给定两个非零的高斯整数 z 和 m,在这个环中,欧几里得除法算法成立,总存在另外两个高斯整数 q 和 r,使得 $z = qm + r$ 且 $N(r) < N(m)$,(这就证明了这个除法算法在代数整数环中一般不存在,使之成立的整环称之为**欧几里得整环**,)然后,戴德金证明恰好与通常整数一样,对任意两个高斯整数 z 和 m 重复使用整数除法,最终会得到一个最大公因数 d,可写成 $d = az + bm$ 的形式,特别地,如果一个不可约高斯整数整除两个高斯整数的乘积 rs,那么它一定整除它的一个因子,因为如果 p 不整除 r,那么 $1 = ap + br$ 或者 $s = aps + brs$,于是 p 整除 s,立即得到惟一分解定理,就像高斯一样,戴德金能确定 Γ_i 的所有素元.

戴德金进一步注意到,若 θ 是下列方程

$$x^2 + x + 1 = 0, \quad x^2 + x + 2 = 0, \quad x^2 + 2 = 0, \quad x^2 - 2 = 0 \quad 或 \quad x^2 - 3 = 0$$

中的任何一个的根的话,在 Γ_θ 中就有相似的除法算法及惟一分解定理,另一方面,在由 $x^2 + 5 = 0$ 的一个根所生成的整环中,算法不成立.在该整环中任何整数 $\omega = x + y\sqrt{-5}$ 的模为 $N(\omega) = x^2 + 5y^2$.戴德金考虑了整数 $a = 2, b = 3, b_1 = -2 + \sqrt{-5}, b_2 = -2 - \sqrt{-5}, d_1 = 1 + \sqrt{-5}, d_2 = 1 - \sqrt{-5}$.利用模的概念证明了上述的整数都是不可约的.更进一步,简单的乘法表明 $ab = d_1d_2$, $b^2 = b_1b_2, ab_1 = d_1{}^2$.由此可得,惟一分解定理不成立.但是,戴德金走得更远:"稍加思考,…… 前述的数是有理整数."[11] 利用广义整除性法则并假定 a 与 b, b_1 与 b_2 是互素的,戴德金指出,将存在整数 α、γ、δ.使 $a = \alpha^2, b = \gamma\delta, b_1 = \gamma^2, b_2 = \delta^2, d_1 = \alpha\gamma, d_2 = \alpha\delta$ 成立.那么,就有 $ab = \alpha^2\gamma\delta = d_1d_2$ 成立,但是这些整数在给定的这个环中并不存在,毕竟,这些原来的整数都是不可约的,寻找这些整数的替代物并保持惟一分解定理成立就成了他创造新观念目标,他认为,理想是比库默尔的理想数更好理解的一个概念.

戴德金坚信:"没有再创造的必要了.比如说,库默尔的理想数,只在已经存在的数系中考虑就已经足够了."[12] 由于库默尔只用理想数定义了整除性,戴德金姑且认为他的"已经存在的数系"为集合 I,即在 Γ_θ 中所有可以被给定的理想数整除的整数,他把这个集合称为**理想**.由于对于任意可以被一个理想整除的数 α 和 β,它们的和 $\alpha + \beta$ 也可以被该数整除.又由于 $\omega\alpha$ 可以被 Γ_θ 中的任意 ω 整除,这种条件对于理想集是必要的.但是任何满足以上条件的集合是可以被 θ 个理想数整除的

集合，戴德金把这些条件看作代数整数环 Γ_θ 上的理想 I 的定义. 戴德金进而定义了**主理想**(α) 为给定整数 α 的所有倍数组成的集合. 如果 α 和 β 都是 Γ_θ 的元素，那么，这个理想包含了所有的形如 $r\alpha + s\beta$ 的整数. 其中 r、s 是 Γ_θ 中的元，记作(α, β).

戴德金的下一个任务是定义理想的整除性，他证明了如果 α 被 β 整除，或者 $\alpha = \mu\beta$，则由 α 生成的主理想包含在由 β 生成的主理想中. 反之，如果 α 的任何一个倍数也是 β 的一个倍数的话，那么 α 可以被 β 整除. 因此，戴德金把这个定义推广到了任意理想上去，一个理想 I 可以被理想 J **整除**，如果 I 中的任何一个元素均在 J 中，一个不同于 Γ_θ 的理想 P 被称之为**素元**，如果它没有除 Γ_θ 和它本身之外的任何因子，即它不含于除 Γ_θ 外的其他理想中. 例如，$\Gamma_{(-5)}$ 中的主理想(2) 尽管它有不可约元素生成，但不是素元，它包含在素理想$(2, 1 + \sqrt{-5})$ 中.

戴德金进一步注意到有一个理想的之积的自然定义，即，IJ 由所有形如 $\alpha\beta$ 乘积的和所组成，这里 α 是 I 的元素，β 是 J 的元素，很明显，IJ 可以被 I 和 J 整除. 但是，为了揭示积和可除性之间的关系，他还得再证明两个定理.

定理 如果理想 C 被理想 I 整除，那么存在惟一的理想 J 使 IJ 与 C 相同.

定理 所有不同于 Γ_θ 的理想或者是素理想，或者可以惟一地由一些素理想的乘积所表示.

正是这两个定理使戴德金能在任何代数整数环上建立惟一分解定理，这样，在形如 $x + y\sqrt{-5}$ 的代数环中，(不存在的) 素因子 α、γ、δ 被理想 $A = (2, 1 + \sqrt{-5})$，$G = (3, 1 + \sqrt{-5})$，$D = (3, 1 - \sqrt{-5})$ 代替，这样主要理想$(a) = (2)$，$(b) = (3)$，$(d_1) = (1 + \sqrt{-5})$，$(d_2) = (1 - \sqrt{-5})$ 就被分解为$(a) = A^2$，$(b) = GD$，$(d_1) = AG$，$(d_2) = AD$. 那么，非惟一的整数因子分解 $ab = d_1d_2$ 就可以被惟一素理想分解$(a)(b) = A^2GD = (d_1)(d_2)$ 代替了.

图 15.2 德国邮票上的戴德金及理想的分解因子.

人物小传

理查德 · 戴德金(1831—1916)(Richard Dedekind)

戴德金(图 15.2) 与高斯一样出生于不伦瑞克镇，并就读于卡丁纽姆学院和哥廷根大学. 尤其是在 1852 年在高斯指导下获得博士学位之后，仍继续在哥廷根和柏林跟着德国最优秀的数学家学习. 在 1858 年，他成为苏黎士工科学校的一位教授，4 年后他从卡罗林纽学院返回不伦瑞克，在这里的工科学校教书. 尽管他有多次去德国主要大学就职的机会，他选择了留在不伦瑞克，在这里他感觉有他从事数学研究的足够的自由. 这是他为狄利克雷的《数论讲义》(1863 年和以后的几版) 做编辑工作使他决定发表他自己在这一学科方面的思想. 这些思想在几年之前他的讲义中就已经形成了. 因此在狄利克雷的第二版著作中(1871) 他提供了一份补充材料，在这份补充材料中他创立了代数整数的理论，并在以后的几版中逐渐扩充了这个理论.

15.2 解代数方程

由于解方程是 19 世纪之前代数学的中心所在，所以在那个世纪新形式的代数所呈现出的主要特征由解代数方程的新方法产生出来就没有什么可惊奇的了. 事实上，群论中的一些主要思想就源

于这些新方法.

15.2.1　割圆方程

拉格朗日已经很详细地研究了低于 5 次的方程的可解性问题,并提出一些解决更高次方程的可能的方法,高斯在《研究》的最后一章中讨论了割圆方程(即形如 $x^n - 1 = 0$) 的解,并把这种解运用到正多边形的构造上.高斯知道这个方程的解是形如

$$\cos\frac{2k\pi}{n} + \mathrm{i}\sin\frac{2k\pi}{n} \quad (k = 0,1,2,\cdots,n-1)$$

的.但在本章中,他的目的是想从代数上确定这些解,由于系数为合数的方程的解很容易从素数次方程的解推出,高斯把他的注意力限制在 n 为素数的情况,又因为 $x^n - 1$ 分解成

$$(x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1),$$

因此方程

$$x^{n-1} + x^{n-2} + \cdots + x + 1 = 0$$

成为了他研究的重点.

高斯解这个$(n-1)$ 次的方程的计划是解一系列辅助方程.它们中每个方程的阶为$(n-1)$ 的一个素因子,每一方程的系数完全由前一方程的根来决定.对于 $n = 17$ 时,由于 $n-1 = 2\cdot2\cdot2\cdot2$,他想构造四个二次方程,当 $n = 73$ 时,他需要构造三个二次方程和两个三次方程.高斯知道 $x^{n-1} + x^{n-2} + \cdots + x + 1$ 的根可表示成任何一个固定根 r 的幂的形式,$r^i(i = 1,2,\cdots,n-1)$.更进一步,他发现如果 g 是任何一个模n 的本原根的话,那么幂 $1,g,g^2,\cdots,g^{n-2}$ 就包含了所有模 n 的非零剩余.由此可以推出方程的$(n-1)$ 各根可以表示成 $r,r^g,r^{g^2},\cdots,r^{g^{n-2}}$,甚至对于任意小于 n 的λ,表示成 $r^\lambda,r^{\lambda g},r^{\lambda g^2},\cdots,r^{\lambda g^{n-2}}$.确定这些辅助方程的方法涉及到了**周期**的构造,即根 r^j 的某些和,它们依次是辅助方程的根.对 $n = 19$ 时的分析作为特定的一个例子把高斯工作的特点呈现了出来.

由于 $n = 19$,所以 $n-1$ 的因子是 3,3,2.高斯首先分别确定了由六项组成的 3 个周期,每个周期都是一个 3 次方程的根,这些周期是通过选择模 19 的本原根,在这里是 2.置 $h = 2^3$,并计算

$$\alpha_i = \sum_{k=0}^{5} r^{ih^k} \quad (i = 1,2,4).$$

用现代术语,方程 $x^{18} + x^{17} + \cdots + x + 1 = 0$ 的 18 个根的置换构成一个由对应关系$r \to r^2$ 确定的循环群 G,这里的 r 是任意一个固定的根.这里的周期是由对应关系 $r \to r^h$ 生成的群G 的子群H之下的不变元的和.因为 $h^6 = 2^{18} \equiv 1(\mathrm{mod}\ 19)$ 这些和包含6项.即是说 H是G 的6阶子群.更进一步,对于 $i = 1,2,4$ 形如 $r \to r^{ih^k}(k = 0,1,\cdots,5)$ 的映射恰好是 H 在群 G 中的 3 个陪集.比如说,因为

$$H = \{r \to r, r^8, r^{64}, r^{512}, r^{4096}, r^{32\,768}\},$$

可以推出

$$\alpha_1 = r + r^8 + r^7 + r^{18} + r^{11} + r^{12},$$

这里的次已约化到模 19.类似的,

$$\alpha_2 = r^2 + r^{16} + r^{14} + r^{17} + r^3 + r^5, \quad \alpha_4 = r^4 + r^{13} + r^9 + r^{15} + r^6 + r^{10}.$$

高斯然后证明 $\alpha_1,\alpha_2,\alpha_4$ 是 3 次方程 $x^3 + x^2 - 6x - 7 = 0$ 的根.

下一步是把这 3 个周期中的每一个又分成 3 个周期,每个由两项组成.这里的 3 个新周期将再满足一个 3 次方程.它们是

$$\beta_i = \sum_{k=0}^{1} r^{im^k} \quad (i = 1,2,4,8,16,13,7,14,9).$$

这里 $m = 2^9$,并在映射 $r \to r^m$ 生成的子群 M 下是不变的.因为 $m^2 \equiv 1 \pmod{19}$,M 是两阶的,并有 9 个陪集对应着给定的 i,比如说,

$$\alpha_1 = \beta_1 + \beta_8 + \beta_7 = (r + r^{18}) + (r^8 + r^{11}) + (r^7 + r^{12}).$$

给定了长度为 2 的新的周期,高斯证明了 β_1,β_8,β_7 是三次方程:$x^3 - \alpha_1 x^2 + (\alpha_2 + \alpha_4)x - 2 - \alpha_2 = 0$ 的所有根.这表明其他的每一个 β_i 都可以表示成 β_1 的多项式.最后,高斯把由两项构成的周期拆成了由这两项分别构成的周期,并证明了,比如说,r 和 r^{18} 是 $x^2 - \beta_1 x + 1 = 0$ 的两个根,剩下的 16 个原方程的根继而就是 r 的简单的幂,或者可以通过解剩下的 8 个相似的 2 次方程而得到.

由于在上面的例子中的 3 次方程正如二次方程一样是利用根式解出的一样,高斯证明 $x^{19} - 1 = 0$ 可根式求解,更一般的结论,应用到任意方程 $x^n - 1 = 0$,只需说明能够找到一系列的次数小于 $n - 1$ 且为素数的方程,它们的解可以确定原方程的解.但是,高斯继续写到:

> 每个人都知道,最杰出的几何学家在寻找高于 4 次的方程一般解法时都没有成功,或者(更精确地把寻找说成)将混合方程约化为单纯方程[指形如 $x^m - A = 0$ 的方程,这个方程在 $A = 1$ 的解知道的情况下,可以通过乘一个 m 次的根而解出来].毫无疑问,这个问题不仅超越了我们现代分析学家的能力,而且提出了它不可能…….尽管如此,我们仍可以确定每一阶方程中都存在无以计数的可以约化为单纯方程的混合方程…….我们相信,如果我们证明了我们的方程都是上述的那种方程的话,几何学家将认为这是非常令人满意的.[13]

高斯然后给出了一个简证,尽管仍有一小缺陷.他证明了在解方程 $x^n - 1 = 0$(n 为素数)时涉及到的辅助方程也可以被化为单纯方程.因此,通过归纳,他证明了这些方程可根式求解.自然,如果 $n - 1$ 是 2 的幂次,所有的辅助方程都是 2 次的,没有特殊证明的必要.但是,在这种情况下,高斯进一步注意到这些解可以通过欧氏几何的技巧构造出来,因为 $x^n - 1 = 0$ 的根可以被看作是一个正 n 边形(在复平面上)的顶点,高斯证明了一旦 $n - 1$ 是 2 的幂,这样的 n 边形是可以构造的.高斯知道,甚至我们今天所知道仅有的这种素数是 3、5、17、257、65 537.事实上,传说高斯对正 17 边形的构造方法的发现帮助高斯决定了他从事数学的生涯.高斯用一个告诫来结束他的著作:"一旦($n - 1$)包含除 2 以外的素数,我们都将面对更高次的方程 …… 我们可严格地证明这些高阶方程无论用什么方式都不可避免,而且也不可以化为低阶的方程 …… 我们发表这个提醒,以免任何尝试找到把圆周等分成除我们的理论提到的数之外的部分(比如 7、11、13、19 等)的几何构造方法的人白白地浪费了他的时间."[14]

有趣的是,尽管高斯已经做出了该断言,但事实上他并没有证明正 n 边形,n = 7、11、13、19 等不能被构造.这个差距于 1837 年被皮埃尔·万泽尔(1814—1848)补上.万泽尔也给出了经典希腊作图问题的解,他证明任何不能导出某阶数为 2 的幂次且系数可作图的不可约多项式方程的作图问题不能用直尺和圆规来完成.比如,因为边长为 a 的倍立方体问题需要不可约三次方程 $x^3 - 2a^3 = 0$ 的解,所以用欧氏工具做出这个解是不可能的.类似地,把一个角 α 三等分的问题需要求解一个用已知的 $a = \sin\alpha$ 来表示 $x = \sin(\alpha/3)$ 的不可约三次方程,即求方程 $4x^3 - 3x + a = 0$ 的解.万泽尔的结果再次证明这种作图方法用欧氏工具是不可能的.

但是回想希腊数学家们用圆锥截线把这两个问题全部解决了.圆锥截线是由奥马·海亚姆和笛卡尔推广出来的.他们给出了解三次方程的具体作图法,相似地,美国数学家 J. 皮尔盼特

(1866—1932) 在 1895 年证明了正 n 边形(n 为素数) 可以用圆锥截线构造出来,当且仅当 $n-1$ 不包含除 2 和 3 之外的素因子. 例如,正 7 边形可以用圆锥截线构造出但 11 边形就不可以. 另一个很重要的构造问题希腊数学家没有解决,即化圆为方,也证明是不可能的. 代数上,这个问题等价于解二次方程 $x^2-\pi=0$,不幸地是这个方程的一个系数是 π,希腊数学家没能找到用欧氏工具画出长为 π 的线段的画法. 到 19 世纪前很长时间人们怀疑 π 不能被表示成任何有理系数的方程的根. 即,它是一个**超越数**而不是一个代数数. 刘维尔在 1844 年首先真正给出了一个超越数:

$$\frac{1}{10}+\frac{1}{10^{2!}}+\frac{1}{10^{3!}}+\cdots+\frac{1}{10^{n!}}+\cdots=0.110\,001\,000\,000\,000\,000\,000\,001\,00\cdots.$$

后来他想尽力证明 e 和 π 都是超越数, 但没有成功, 刘维尔的一位被保护者 C. 埃尔米特 (1822—1901),在 1873 年最终证明了 e 是超越数,在埃尔米特思想的基础上,费尔迪南德 · 林德曼 (1852—1939) 在 9 年之后证明了 π 是超越数. 由此马上可以推出用欧几里得工具化圆为方是不可能的.

15.2.2 置换理论

非常明显,高斯确信高于 4 次的一般方程不可用根式求解. 回想拉格朗日曾经试着考虑根的置换而致力于找到其解. 因此,为了详细考虑高阶方程问题,弄明白置换的理论是必需的. 这个方面的实质性的工作在 19 世纪初期被柯西(1789—1857) 完成了.

一直到柯西的时代,置换一词通常是指某些物体(比如字母) 的排列. 是柯西首先意识到了从某一种排列变换到另一种排列的重要性. 他用了**替换**一词来表示这个动作,就是我们现在所说的"置换",是指元素的有限集合上的一个一对一的函数,从 1815 年最初的研究之后近 30 年中,在关于这个问题的一系列论文中柯西用'替换' 和'置换' 这两个词来不加区分地指代上述的函数. 为了防止混淆,我们在此统一地用现代意义的后者.

除了注意到置换的函数性质方面之外,柯西还用了单个字母 S 来表示一个给定的置换,并且定义了两个这种置换的乘积,记为 ST,它是指一种置换,这种置换先把 S 用于一个固定的"排列",而后再把所得结果作用于 T 变成最后的"排列". 他称使给定的排列不变的置换为恒等置换,并注意到给定的变换 S 的幂: $S, S^2, S^3\cdots$ 一定会出现一个与 S 相同的变换,并定义变换 S 的**阶**是使 S^n 与恒等变换相同的最小的 $n(n\neq 1)$,柯西还在字母表 $a_1, a_2, \cdots, a_n$ 上定义了他称之为**循环**置换的置换,是指把 a_1 变到 a_2, a_2 变到 $a_3, \cdots, a_{n-1}$ 变到 a_n, a_n 变到 a_1 的置换. 在 1844 年,柯西用记号$(a_1, a_2, \cdots, a_n)$ 表示这个置换,同时他也明确地定义了置换 S 的逆,用符号 S^{-1} 来表示,1 是恒等置换. 更进一步,在给定的任何一个含有 n 个字母的集合上,他定义了**共轭置换系**为由最初的置换经过所有可能的乘法而得到的置换的集合. 也就是我们今天所称的由给定集合生成的**子群**. 最后他证明了这个系的阶(系中元素的个数) 总是整除整个的 n 个字母的置换系的阶 $n!$.

15.2.3 5 次方程的不可解性

一般的五次方程不能用根式求解的第一个证明出现在意大利人鲍罗 · 鲁菲尼(1765—1822) 在 1798 年发表一篇私下印刷的论文上的. 它的证明在当时没有人可以理解. 然而 19 世纪 20 年代中期,阿贝尔(1802—1829) 最终给出了这个方法不可能性的完整的证明.

阿贝尔的不可解性证明涉及到关于方程的根的置换的结果. [16] 但是,我们注意到在证明了他的不可解性结论之后,阿贝尔尝试着解决下列的问题:"1) 对于给定的阶,找到所有可代数解的方程; 2) 确定一个给定的方程是不是代数可解的."[17] 尽管在他的余生中没能完全解决上述问题中的任何

一个，但在一些特殊的类型的方程上他的确有了很大的进展. 在 1829 年发表在克雷尔的《杂志》上的一篇论文中，阿贝尔推广了高斯关于方程 $x^n - 1 = 0$ 的解法. 对这个方程而言，它的根可以用其中的一个根的幂来表示. 阿贝尔证明了“如果一个方程任何阶的根都可以用某一个根(命名为 x)来有理表示. 而且，如果对任两个根 θx 和 $\theta_1 x$[θ 和 θ_1 是有理函数]，我们有 $\theta\theta_1 x = \theta_1\theta x$ 的话，那么该方程就是代数可解的.”[18] 他证明这个结果是通过证明在这种情况下，与在割圆方程的情形那样，总可以把这个解约化到素数阶的辅助方程上去. 就是由于这个结果，我们把今天所说的交换群通常也叫阿贝尔群.



图 15.3 挪威邮票上的阿贝尔.

人物小传

尼尔斯 · 亨利克 · 阿贝尔(1802—1829)(Niels Henrik Abel)

阿贝尔，出生在挪威斯塔万格附近，非常可惜的是他的一生非常短暂. 在奥斯陆的教会学校上学的时候，他的数学老师发现了他在数学方面的天赋，并鼓励他读那些在大学可以得到的各种前沿的论文. 他对五阶方程变得非常感兴趣，他相信他能用根求解. 事实上，他也朝这个方向做了，因为在挪威没有人能懂他的证明，他请人把这篇论文寄到了丹麦. 在文章发表之前，阿贝尔被要求提供几个例子. 在寻找这些例子的时候，他意识到他的证明是不正确的. 尽管他后来转向了其他方向的研究，特别是椭圆函数理论的研究，在他在奥斯陆上大学的几年里，阿贝尔仍然在可解性问题上继续向下走，直到他证明该问题的不可解性使自己满意为止. 在 1824 年，阿贝尔自己出钱印了一个关于这个结论的小册子. 但是这种为了省钱的简短使许多数学家不能理解它. 因此两年以后，在他周游欧洲去拜访不同的数学家，并为他的科学生涯更好做准备期间，他把扩充的版本发表在德国一本新的数学杂志《纯粹和应用数学的杂志》的第一卷上. 这本杂志由克雷尔编辑，很快他们就成为了最亲密的朋友. 当阿贝尔于 1827 年返回挪威的时候，发现他已经没有任何的职位了. 大学里惟一的一个数学教授头衔刚刚授予了他中学的数学老师. 阿贝尔通过当助教和给别人代课而艰难地谋生，同时也准备好了新的数学论文，但是在 1829 年患上肺结核之后一直没有康复，他于四月份去世，也就是克雷尔写信告诉他在柏林已经找到新的职位的两天之前.[15]

15.2.4 伽罗瓦的工作

尽管阿贝尔没能完成他的研究项目，但是大部分都被另一位早逝的天才埃瓦里斯特 · 伽罗瓦(1811—1832)完成. 伽罗瓦关于代数方程根式求解性问题的思想，初步形成于他在 1831 年呈交给法国科学院的手稿中. 在这份手稿中，他是以弄明白有理性思想为开端的. 既然方程的系数在某个域之中，不妨设为普通的有理数域，称一个方程用根式可解是说每个根可以对域中的数用基本的四则运算和开根方来表示. 但是逐步解方程显得更方便，就如高斯在割圆方程时所做的一样. 因此，比方说，一旦解出了方程 $x^n = \alpha$，下一步，系数为$\sqrt[n]{a}$，$r\sqrt[n]{a}$，$r^2\sqrt[n]{a}$，…，的解就是可用的了，这里 r 是一个 n 次单位根. 伽罗瓦注意到这些量添加到原来的域中，用这些量以及原来域中的量的四则运算表示的量都可以看作**有理的**(用现代术语说，即是开始于一个特殊的数域，通过添加某些不在原来域中的数而构造的扩域). 当然，“方程的性质和难点可能完全不同，这依赖于添加什么样的量.”[19] 伽罗瓦也介绍了他在置换方面的想法，与柯西一样用了含混的语言，使用了词语“群”，同样是不严格的意义，有时指在复合运算下封闭的置换的集合，有时又指通过应用某个置换而确定的一些排列的集合.

伽罗瓦的主要结果表达如下：

命题 I　假定一个方程的根是给定的 m 个数 $a,b,c,\cdots$(伽罗瓦默认该方程是不可约的，且它的所有根都是不同的)，则将有字母 $a,b,c,\cdots$ 的一个置换群满足下述性质：

(1) 每个在群的替换下不变的根的函数是有理的；

(2) 相反地，这些根的每一个有理函数在这些置换下都是不变的.

伽罗瓦把这个置换的群称之为**方程的群**，在现代习惯中，通常把根的每一个置换扩充为整个域的自同构，这个域是通过把方程的根加到原来的系数域而得到的. 这样，伽罗瓦的结论就成为存在新数域中一组自同构，其不变元正是原数域中的元素(即有理的已知元素). 伽罗瓦除了扼要地证明了这个结论之外，还给出了高斯的割圆方程多项式的例子$\frac{x^n-1}{x-1}$(n 是素数). 在这种情况下，假设 r 是一个根，g 是模 n 的原根，根就可以表示为如下形式：

$$a = r, b = r^g, c = r^{g^2}, \cdots.$$

这个方程的群是由循环($abc\cdots k$)产生出的 $n-1$ 个置换的循环群. 另一方面，阶为 n 的一般方程(即系数为文字)的群是所有的 $n!$ 个由 n 个字母的置换构成的群.

主要定理确立之后，伽罗瓦开始探求在方程可解性问题上的应用. 他的下一个命题说明了当添加辅助方程的一个或所有的根到原数域(或者初始方程)时将会怎样. 因为任何一个使新域不变的自同构，一定也使原来的域不变，所以新域上的方程的群 H 是原来域上的群 G 的子群. 事实上，G 可以分解为 $G = H + HS + HS' + \cdots$ 或者 $G = H + TH + T'H + \cdots$ 这里 S,S',T,T' 是适当选择的置换. 伽罗瓦在他给谢瓦里耶的信中解释了整个步骤，并指出通常情况下这两种分解不同时存在. 但是，当这两种情况同时出现时，往往是某个辅助方程的所有根都添加进来了. 他称这种分解为**合适的**. 如果一个方程的群有一个合适的分解，那么它就可以划分成 m 个 n 元置换群，这样就可能通过两个方程来解所给的方程. 这两个方程一个有 m 元置换群，另一个有 n 元置换群. 用现代术语，即当子群 H 是**正规**的时合适的分解才存在，即右陪集$\{HS\}$与左陪集$\{TH\}$相同. 在这种情况下，方程的可解性问题就归结为另两个方程的可解性上了，这两个方程的群的阶比原来的小.

高斯很早就已经证明多项式 x^p-1(p 为素数)的根可根式解，由此推知：若在原始的域中有 n 次单位根的话，那么方程 $x^p-\alpha$ 的某个根的添加就是所有根的添加. 若 G 是一个方程的群，那么添加引出群 G 的一个正规子群 H，H 在 G 中的指数(G 的阶除以 H 的阶得的商数)是 p. 伽罗瓦同时证明了反过来的情况：如果一个方程的群 G 有一个指数为 p 的正规子群，那么，可以找到原来域(假定 p 次单位根在原来域中)中的一个元素 α，$\sqrt[p]{\alpha}$ 的添加把方程的群约化成 H. 伽罗瓦在手稿和信中推断说：一个方程是根式可解的，只要寻找正规子群的过程可以一直继续到所有子群的指数均为质数. 伽罗瓦给出了一般情况下四阶方程的详细过程. 证明了方程的阶为 24 的群有一个 12 阶的正规子群，依次的该子群包含一个阶为 4 的子群，阶为 4 的子群又包含一个阶为 2 的子群，阶为 2 的子群又包含单位元. 由此可得，先添加一个二次根，之后是一个三次根，而后再添加两个二次根就可以得到方程的解. 伽罗瓦指出四次方程的标准解法用的恰好是这些步骤.



图 15.4　法国邮票上的伽罗瓦.

人物小传

埃瓦里斯特·伽罗瓦(1811—1832)(Evariste Galois)

伽罗瓦的悲惨而短暂的一生早就成为传记小说的题材.小说中包括某些假想的东西:因为他的政治观点太激进,政府策划了他在决斗中的死亡.但是,已知的事实并不支持这个观点.伽罗瓦出生于赖纳镇,一个距巴黎不远的城镇,1815年他的父亲被选为该镇的镇长.他在赖纳镇上小学时曾获得过各种各样的成就,尤其是发现他的数学天赋之后.尽管他在18岁之前发表过一个短篇论文并同时呈交给法国科学院一篇关于素数次方程的可解性问题的论文,但他两次都没能通过综合工科学校的入学考试.第一次大概是因为他没能掌握好基础知识,第二次可能是因为他父亲于几天前刚刚由于反动牧师的诽谤而自寻短见.因而伽罗瓦不得不进入师范学校学习.这里的老师把学生们锁在学校里以防止参加1830年7月革命之前的那些政治运动.当他在12月份的一封信里攻击他的指导老师比起"自由"来更偏爱"正统"之后,伽罗瓦被开除出了学校.随之参加了一个共和党色彩很重的组织:国民卫队.由于这个组织对"资产阶级"路易·菲力普国王的王位有一定的威胁而很快被解散了.虽然卷入了复杂的政治运动,伽罗瓦仍继续着他的数学研究.1831年他把一篇方程可解性问题的修改论文投送给了研究院.六个月之后审稿人把这篇文章退了回来,因为读不懂他的这个证明,并建议伽罗瓦把论文完善、简明化之后再重新投稿.

但是,那时伽罗瓦已经被捕过两次了.第一次由于威胁要国王的命,第二次由于他穿了已被解散组织的国民卫队的制服.这一次他被宣布有罪,并判入狱六个月.在此期间,他由于科学院不能理解他的论文而憎恨他们,以至于在他没有公开出版的作品的序言中尖刻的讽刺了法国的"官方科学家".但是在他的文章发表之前,伽罗瓦又卷入了"一个邪恶女人和她的两个骗局"[22]之中.尽管具体的事情还没弄清楚,伽罗瓦就被迫(挑选)去参加一场决斗.伽罗瓦被击中致死,当时还有五个月到他21岁生日.在他去参加决斗的前一天晚上,他预感到了死亡.于是给他的朋友A.谢瓦里耶写了一封信,详细的阐述并注解了他以前的部分论文,并以下述的话结束:我一生中常常敢于提出一些我自己不太确定的主张,但是我上述所写的早在一年多以前就在我头脑中很清楚了,发表自己还没有完全证明的理论不是我兴趣所在,请求雅可比或者高斯评判这些定理是否重要而不是是否正确.我想,在此之后就会有人发现理解所有这些混乱的东西将会非常有用.

15.2.5 约当和群的置换理论

伽罗瓦死后,他的手稿一直没有发表,直到1846年才由刘维尔出版在他《数学杂志》中,在之后的几年中,一些数学家在大学的讲课中引用了伽罗瓦资料或者发表了对他工作的评论.但是直到1866年伽罗瓦的理论才出版在保罗·舍纳特(1827—1898)的《代数教程》第三版的教材中.四年之后,卡米尔·约当(1838—1922)出版了他的纪念文章《论置换和代数方程》,其中包含一些改进了的伽罗瓦理论.

群论的许多现代观念首先出现在约当的教材以及他前十年间发表的、主要内容已被编进该教材的一些论文中(尽管仍然是以置换群为背景).约当定义了**群**是有限集的一组置换,并且其中任何两个置换的乘积(复合)仍在其中.之后他证明了每个组都包含一个单位元1,且对任一置换 a 有另一置换 a^{-1} 使 $aa^{-1}=1$.约当定义了置换 b 对置换 a 的**变换**为置换 $b^{-1}ab$,b 对群 $A=\{a_1,a_2,\cdots,a_n\}$ 的变换为 $B=\{b^{-1}a_1b,\ b^{-1}a_2b,\cdots,b^{-1}a_nb\}$,如果 B 与 A 一样,A 就称为与 B **可置换**.尽管约当没有明确定义群的正规子群,但他定义了**单群**是不包含任何除单位元之外的对自己的每个元都可置换的子群的群.对于一个非单群 G,一定存在一个**复合序列**,即一个群列 $H_0,H_1,H_2,\cdots,\{1\}$,其中后

一个群包含在前一个群中,并且对群的每个元素都是可变换的(即正规的),而且没有其他的群可以插入这个序列之中.约当进一步证明,如果 G 的阶是 n,且子群的阶依次为 $\frac{n}{\lambda},\frac{n}{\lambda\mu},\frac{n}{\lambda\mu\nu},\cdots$,那么整数 $\lambda,\mu,\nu,\cdots$ 对阶来说是惟一的,即其他的任何序列也有相同的**复合因子**.[25]

约当然后用群论的语言重述了伽罗瓦结论,他把可作为根式可解的方程定义为**可解群**.因此,可解群是一种包含某个复合序列的群,该数列中的复合因子皆为素因子.[26] 因为交换群往往包含合成素因子,约当证明了**阿贝尔方程**,即"其群仅含与它们自身可交换的置换",总是根式可解的.另一方面,由于 n 个字母的交错群(阶为 $n!/2$) 当 $n>4$ 时是单的,由此很快可得 n 次的一般方程不是根式可解的.约当解释清楚了伽罗瓦的工作,显然置换群的理论与方程的可解性是紧密相连的了.

15.3　群和域 —— 结构研究的开始

群论的某些概念在 19 世纪早期的数论和方程的根式可解性的发展中就已隐含了.高斯在上述两个领域中起了很重要的作用.他在二次型理论方面的论述也是同样重要的.正是这个结论使上述的两种思想最终成为群的抽象理论的一部分.

15.3.1　高斯和二次型

高斯在他《数论研究》的第 5 章中讨论了二次型的理论,即关于两个变元 x、y 形如 $ax^2+2bxy+cy^2$ 函数理论.这里 a、b、c 是整数.高斯讨论的最初目的是确定一个给定的整数能否表为特殊的形式.作为解决该问题的工具,他定义了两个形式的等价.形式 $f=ax^2+2bxy+cy^2$ **等价**于形式 $f'=a'x'^2+2b'x'y'+cy'^2$ 是指,如果存在线性变换 $x=\alpha x'+\beta y'$, $y=\gamma x'+\delta y'$,且 $\alpha\delta-\beta\gamma=1$,把 f 转变成 f'.通过简单的计算可以证明任何两个等价二次型具有相同判别式 b^2-ac.另一方面,具有相同判别式的形式不一定等价.高斯证明了对于任意给定的判别式的值 D 有无穷多类等价二次型.特别的,有一个特殊类 —— 主类,是由等价于 x^2-Dy^2 的形式所组成的类.

为了研究这些类,高斯给出了二次型的一种复合规则.换句话说,对于给定的有相同判别式的形式 f,f'.高斯定义了由 f、f' 复合所得的新形 F(记为 $F=f+f'$).他具有一些所需的性质.首先高斯证明如果 f 和 g 等价且 f' 和 g' 也等价,那么 $f+f'$ 等价于 $g+g'$.因此复合运算是类上的一个运算.高斯下一步证明了复合运算是可交换和可结合的.最后,高斯证明"如果类 K 与主类复合,结果将是类 K 本身;" 对任意类 K,存在类 L(K 的相反) 使得这个两个类的复合为主类;"对于给定的有相同判别式的两个类 K、L,我们总能找到有相同判别式的类 M 使 L 是 M 和 K 的复合".给定的这种复合满足基本的加法性质,高斯指出:"用加号表示类的复合,用等号表示类的相同是非常方便的."[27]

用加号作为运算的符号,高斯用 $2C$ 来表示类 C 与它自己的复合,用 $3C$ 表示 C 与 $2C$ 的复合,依次类推,高斯证明对于任意的类 C,存在一个与主类相等的最小的倍数 mC,且如果类的总个数为 n 的话,m 是 n 的因子.很自然,这个结论使他想起了《数论研究》的早期材料."上述理论的证明与剩余类的乘幂的证明非常相似.事实上,类的复合理论在每一方面都与前面处理过的课题紧密相联".[28] 因此,他不加证明地断言了许多其他来自这个相似性的结论,这些都可以用现在的阿贝尔群的理论的词语表示出来.

15.3.2　克罗内克和阿贝尔群的结构

尽管高斯意识到了上述两种处理方法的相似之处,但也没有尝试去发展群的抽象理论.这方面

的发展经历了很多年，19世纪40年代中期，库默尔在研究他的理想复数的时候，注意到它在许多方面与高斯关于型的理论非常相似．特别的，库默尔定义了理想复数的等价性，他把它们划分成一些类，这些类的性质与高斯二次型的类的性质相似．但最终是列奥波德·克罗内克(1823—1891)，库默尔的一个学生，看出了抽象理论可以从这些相似之中提取出来．

在1870年的一篇论文中，他发展了库默尔的理想复数中的数类的性质，克罗内克回忆了高斯的二次型的论述：

> 高斯的方法所依赖的非常简单的原则不仅在那个给定的领域中适用，而且在其他方面也常常用到，尤其是数论的基础部分．这种情况表明这些原则属于一个更加广泛和抽象概念的领域，这一点很容易让人信服．因此，把它们的发展从那些不重要的约束中解放出来就非常必要．这样一个人就可以把自己从不同的情形下必须重复相同的证明之中解放出来！这种优点还在其自身发展中显现出来，如果是以最一般的可接受的方式给出的话，其表述变得简洁，因为最重要性质就会突现出来．[29]

于是，克罗内克就开始发展这些简单的原则："设 $\theta', \theta'', \theta''', \cdots$ 是有限个元素，对于任何一对元素就有第三个元素依某一确定的程序与之相关联，克罗内克进一步要求这种关联性是可交换的和可结合的．他起初把关联性记为 $f(\theta', \theta'') = \theta'''$，后来又写成 $\theta' \cdot \theta'' = \theta'''$，还要求如果 $\theta'' \neq \theta'''$，则 $\theta'\theta'' \neq \theta'\theta'''$．由于假定元素个数有限克罗内克进一步推演出存在单位1，且对任何元 θ，存在最小的幂 n_θ 使 $\theta^{n_\theta} = 1$ 成立．

最后，克罗内克发展了有限阿贝尔群的基本定理，即存在由元素 $\theta_1, \theta_2, \cdots, \theta_m$ 组成的群中的一个元素 θ 可惟一的表成形如 $\theta_1^{h_1}\theta_2^{h_2}\cdots\theta_m^{h_m}$ 的乘积的形式，在这里对每个 i 有 $0 \leqslant h_i < n_{\theta_i}$．进一步，$\theta_i$ 可以被排成一列，使每一个 n_θ 可被他的后继整除且这些数的乘积恰好是群中元素的个数．随着抽象定理的证明，克罗内克以不同的方式解释了这些元素．指出多种情形下的相似结果都已用其他的结果在前面证明过了．

人物小传	列奥波德·克罗内克(1823—1891)(Leopold Kronecker)
	克罗内克，为了尽可能的得到最好的数学教育，在1845年于柏林获得博士学位之前先后在柏林大学、波恩大学和布雷斯劳大学学习，之后的几年他致力于管理家族商业，经济上完全独立．由于以数学研究作为自己的嗜好，在1861年被选送到柏林研究院，并允许在大学里讲课．1880年，他接管了克雷尔的《数学杂志》的编辑工作．三年之后，由于库默尔退休了，克罗内克成为柏林的数学教授，与魏尔斯特拉斯一起指导着这里的极有影响的数学讨论班．

15.3.3　凯莱和群的定义

非常有趣的是克罗内克即没有为他定义的系统命名，也没有用产生于伽罗瓦理论的置换群进行解释．克罗内克也很可能不知道16年之前亚瑟·凯莱(1821—1895)就已经在置换群概念的基础上发展了相似的抽象理论．

凯莱在他的《群论》中指出置换群的想法归因于伽罗瓦，并继而很快推广到了定义在某个量的集合上的运算或函数的集合上．他用符号1来表示使所有量不变的函数，并指出对于函数，存在一个明确的且可结合的复合概念，但它一般不可交换．但是后来凯莱把这种基本思想从具体的运算概念中提炼了出来，并定义了**群**是"不同符号1、α、β、$\cdots$ 的集合．这样任何两个元的乘积(无论按什么

顺序）或者它们中任何一个元与自身的乘积，都属于这个集合 …….由此可得，如果整个群与其中一个符号相乘，左乘或右乘，结果只不过是把这个群重新生成了一次.”[30] 用现在的观点，我们可以看出凯莱忽视了定义一个很重要的部分.正因如此，凯莱的最后一个叙述为什么成立就显得很不清楚.尽管如此，凯莱基本上想到了有限集上的置换群，从而假定集合的符号是有限的以及乘积是可结合的，且每一个符号都有一个逆的另一个符号，它与前一个符号的乘积为 1.

为了研究不同的抽象群，凯莱介绍了群表：

	1	α	β	$\cdots$
1	1	α	β	$\cdots$
α	α	α^2	$\beta\alpha$	$\cdots$
β	β	$\alpha\beta$	β^2	$\cdots$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$

正如他的定义中所声明的那样，上表中的行和列都包含了群的所有符号.更进一步的，如果群中有 n 元素，凯莱指出每个符号都满足方程 $\theta^n = 1$.

凯莱通过一个熟悉的论证说明了如果 n 是素数，群就必须是下面的形式：$1,\alpha,\alpha^2,\cdots,\alpha^{n-1}$.如果 n 不是素数，就有其他的可能，特别的，他展示了 4 个元素的两个可能群的群表，以及 6 个元素的两个可能群的群表.之后在 1859 年的一篇论文中，他通过列出元素，定义关系以及每个元等于 1 的最小幂次（指数）的方法描述了所有的阶为 8 的 5 个群.比如，一个群包含元素 $1,\alpha,\beta,\beta\alpha,\gamma,\gamma\alpha,\gamma\beta,\gamma\alpha\beta$，并具有关系 $\alpha^2=1,\beta^2=1,\gamma^2=1,\alpha\beta=\beta\alpha,\alpha\gamma=\gamma\alpha,\beta\gamma=\gamma\beta$ 群中除了单位元以外元素的阶都为 2.

尽管凯莱在 1860 年为《英国大百科全书》写了一篇解释“群”和别的术语的文章，在之后的几年中欧洲大陆的数学家没有太多地注意这个（近乎）抽象的定义，也没有人注意到可交换条件的定义定义了相同的结构.但是在 1878 年凯莱发表了关于同一主题上的四篇新论文，在论文中他重述了他的定义和 1854 年的结论.特别的，他写道“群是由符号以及所满足的运算规律来定义的.”[31] 进而，“尽管上述所提的（群的）理论是一般的，包含了置换群作为特例，但是，寻找所有阶为 n 的群的一般问题事实上与看起来更特殊一点的下述的问题是等价的，即寻找所有阶为 n 的由 n 个字母的替换形成的群.”[32] 凯莱把我们今天把它称之为凯莱定理的结论看作是近乎自然的，这只需注意到群中的每个元素都可以看成是通过群运算作用于群的所有元，这种运算诱导了群元的置换.但是凯莱指出：“这并不表示处理一般问题的最好或最简单的方式是把它看作一个置换的问题.很明显，较好的方法是考虑一般的问题并由它导出置换群.”[3] 与克罗内克一样，凯莱意识到群论中的问题往往通过抽象群的考虑而获得解决，而不是具体的实现.事实上，往往只有通过抽象处理才能取得进步.

人物小传

亚瑟·凯莱（1821—1895）（Arthur Cayley）

凯莱在剑桥的三一学院学习数学，并以第三名的成绩毕业.由于找不到合适的教学工作，他决定当律师，于 1849 年进入律师界.尽管他在法律上的工作非常熟练，他把这些看作是提供收入的方式并经常为数学保留一部分的时间.事实上，在他当辩护律师的 14 年间写了将近 300 篇数学论文.在 1863 年，他被选为剑桥新创建的萨多瑞数学教授，尽管他的收入减少了，但他仍非常愿意得到这个职位.

萨多瑞教授的职责是教“纯数学原理”并且“他本人要促进这门学科的发展”.关于第一职责，凯莱做得并不很成功，他在大学里讲课往往只能吸引很少的同学，部分原因是由于他常常讲他最新的研究.另一方面，他对数学的贡献是巨大的，撰写了不同领域的论文近 1000 篇.除此之外，他为其他人的上百篇论文审稿，并热衷于鼓励年轻人开始他们的研究.

15.3.4 群概念的公理化

1878年之后,人们很快意识到了克罗内克和凯莱的定义可以联合引出一个简单抽象群的概念.因此在1879年乔治·弗罗贝尼乌斯(1849—1917)和路德维希·斯蒂克尔伯格(1850—1936)合写的一篇论文这样开头:“对给定的模互素且不同余的数类”,却“不用这些元素的特殊形式”来处理这些类.他们做了如下定义:“如果集合中任意两个元的乘积仍在这集合中的话,这个集合就构成一个(有限)群.”[34] 三年之后,欧根·内托(1846—1919)在教材中详细地,几乎逐字地叙述了克罗内克在1870年的定义.尽管他只涉及到了置换群.

但是,在1822年还有另外两篇出版物可以说是完整的群概念公理化的标志.瓦尔特·迪克(1856—1934)首先出版了他的“群论研究”,在其中他这样陈述基本的问题:“为定义离散运算的群,该群作用在一个特定的对象上,我们忽略单个运算的特殊形式,仅把它们看作由形式群的那些基本性质所推出.”[35] 尽管提到了结合性和逆的性质,迪克在定义群的时候并没有给出这些,而是说明了如何利用生成元和关系来构造一个群.即他首先给了有限个运算,$A_1,A_2,\cdots,A_m$,而后通过考虑这些元素的逆以及他们的幂的所有可能的乘积,建立这些元上的最一般的群 G.这就是我们今天所说的$\{A_i\}$上的**自由群**,自动满足现代的群的公理.进一步迪克通过假定各种形如 $F(A_1,A_2,\cdots,A_m)=1$ 的关系而特殊化为其他群.事实上,他证明如果群 $\overline{G}$ 是由满足给定关系的运算$\overline{A_1},\overline{A_2},\cdots,\overline{A_m}$ 形成的话,那么“G 中所有那些在 $\overline{G}$ 中等于单位元的无穷多个运算构成一个(子)群 H,并与群 G 的所有的运算 $S,S',\cdots$ 可交换.”[36] 而后,迪克证明了映射 $A_i\to\overline{A_i}$ 定义了他称之为从G到$\overline{G}$上的一个**同构**.用现代的术语说就是迪克证明了子群 H 在 G 中是正规的,并且 $\overline{G}$ 与因子群 G/H 同构.

当年的第二篇论文是由海因里希·韦伯(1842—1913)关于二次型的文章.这篇论文首先给出了有限群的完整的公理化描述而未涉及组成它的元素的特性:

> 由 h 个任意东西$\theta_1,\theta_2,\cdots,\theta_h$ 作为元素组成的集合G被称为h阶的群,如果它满足以下条件:
>
> 1.通过一些被称之为复合或相乘的规则,可以由任意两个元导出 G 中的一个新元.用符号表示为:$\theta_r\theta_s=\theta_t$;
>
> 2.总有$(\theta_r\theta_s)\theta_t=\theta_r(\theta_s\theta_t)=\theta_r\theta_s\theta_t$ 成立;
>
> 3.若 $\theta\theta_r=\theta\theta_s$,由 $\theta_r\theta=\theta_s\theta$ 可得$\theta_r=\theta_s$.[37]

通过给定的公理和群的有限性,韦伯推出存在惟一的单位元,并且对于任意一个元素,存在惟一的逆元.他进一步定义如果一个群的乘积是可交换的话就是**阿贝尔群**,而后用了与克罗内克本质上相同的方法证明了阿贝尔群的基本定理.

尽管在以后的几年中,对抽象群概念的应用变得更加普通,但是直到1893年韦伯才发表了包含无限群的定义.他重复了1882年时他的三个条件,并证明如果这个群是有限的,这些条件能确保如果知道群的三个元A,B,C中的任意两个则方程$AB=C$就有惟一的解,另一方面,这个结论对无限群不成立.在那种情况下,我们必须假定方程 $AB=C$ 存在惟一解作为第四个公理,这条公理蕴含了群的每个元都有惟一的逆和惟一的单位元.

在定义了群同构的现代概念之后,韦伯阐明了他抽象方法的基础:“我们可以把所有同构的群当作一个群类,它们本身又是一个群,它的元素是那些把单个的同构群的元素当作一般概念所得的一般特征.单个的同构群被看作一般概念的不同的代表,选哪个代表来研究群的性质都没有关

系."[38] 韦伯构造了许多群的例子,包括平面上向量的加法群,有限集上的置换群,模 m 的剩余类加群,与 m 互素的模 m 的剩余类乘法群,高斯的复合下判别式给定的二次型类的群.随着这些材料的发表,以及 1895 年在韦伯的《代数读本》的再次出现,群的概念已经成为了数学主流中的一部分.

15.3.5　域的概念

域论的历史比群论的要好讲多了,域的概念隐含在 1830 年左右在高斯的工作中.回想伽罗瓦讨论了有理量是什么意思以及如何把一个新的元素添加到给定的有理量的集中.对伽罗瓦来说,有理数域 Q 和由一个超越量或由给定方程的一个根所生成的扩域 $Q(\alpha)$ 直觉上是清楚的,在这里没有必要定义这个概念.克罗内克于 19 世纪 50 年代开始在实际构造这些域时更具体了.克罗内克相信代数和分析可以建立更加严密的基础,通过把所有的概念都建立在始于整数构造的基础上:"上帝自己创造了整数,剩下的就是人的工作了."[39] 因此他觉得像$\sqrt{2}$ 那样的无理数是没有意义的,除非可以找到一个由整数构造它们的确定的方法.而后,他在域的意义下想找到一种构造有理数或其他已确定的域的扩域的方法,它不依赖于无理数的先验存在性.

克罗内克从某些元 $R', R'', \cdots$ 决定的有理域的概念开始,这个域包含所有 $R', R'', \cdots$ 的整系数有理函数.因而他假定了整数以及有理数的存在性.而后他可以解决他的把$\sqrt{2}$ 添加到某有理域内的问题了.在该域中通过考察有理数系数的多项式被 x^2-2 所除的余式可知,x^2-2 没有根.因为两个有相同余式的多项式看作是相等的,可直接在余式集上定义基本运算,因而构成了一个新的有理域.这个构造的另一看法是:把新的有理域看作是包含新元 α 和 α 的所有有理函数,再加上 α^2 总可用 2 来代替这一条件.

在 19 世纪 50 年代,与添加的过程相比,戴德金开始更加关心元素集本身.回想一下,戴德金对代数整数以及可以被表示成代数方程的根的复数的算术非常感兴趣.因而在他 1871 年出版在狄利克雷的《讲义》第二版的补充教材中给出了以下定义:"一个实数或者复数 α 的系 A 叫做**域**,如果对其中的每一对数,他们的和、差、积、商仍都在 A 中."[40](他指出 0 在任何除法中不能做除数,因而域至少包含一个非零元.)当然,最小的这种系就是有理数域,他包含在任何其他的域中,而最大的域为复数域,它包含每个域.因而与克罗内克不同,对于戴德金,一个域中总发生在添加代数元的复数域中.事实上给定任一复数集 K,戴德金定义域 $Q(K)$ 为包含 K 的最小的域.

对戴德金和克罗内克来说任何域都包含有理数域,但谁都没有试着把它的定义推广到其他类型的域中去,尽管早在 1830 年伽罗瓦就发表了简短的本质是描述有限域的论文,伽罗瓦在文中的目的是推广高斯解同余方程 $x^2\equiv a(\bmod p)$ 的想法.伽罗瓦问当无解时会怎样,正像我们为方程 $x^2+1=0$ 创造了解 i 一样,对同余方程 $F(x)\equiv 0(\bmod p)$ 指定一个解 i(这里 $F(x)$ 是 n 阶的并且模 p 的余数本身不是解).伽罗瓦考虑了 p^n 个表达式$a_0+a_1i+a_2i^2+\cdots+a_{n-1}i^{n-1}$,这里$0\leqslant a_j<p$,并注意到这些表达到可以以明显的方式加、减、乘和除.伽罗瓦进一步证明如果 α 是集合中的任一个非零元,α 的某个最小的幂$\alpha^n=1$.用与高斯的模 p 剩余相似的讨论,可证所有这样的元满足 $\alpha^{p^{n-1}}\equiv 1$ 并且有一个本原根 β 使每一个非零元都是β 的幂.伽罗瓦通过证明对任意的素数次幂 p^n,总可以找到一个不可约的 n 阶同余式,它的一个根生成了今天所说的 p^n 阶的伽罗瓦域并以此来结束他的论文.伽罗瓦评注说,找这种多项式的最简单的方法就是通过试验和差错.作为一个例子,他说明了 x^3-2 是模 7 不可约的.因此集合$\{a_0+a_1i+a_2i^2\}$ 的元素构成了阶为 7^3 的域,这里 i 是多项式的 0 点,且当$j=0,1,2$ 时 $0\leqslant a_j<7$.

最终是韦伯把戴德金 - 克罗内克型的域定义和伽罗瓦的有限系的定义结合起来,在 1893 年的

同一篇论文中给出了域的抽象定义.事实上,在他的定义中用了群的观念.**域**就是一个有两种运算:加法和乘法的集合.在第一种运算下是一个交换群.在第二种运算下集合的非零元组成一个交换群,这两种运算通过下面的法则相联系:

$$a(-b)=-ab;a(b+c)=ab+ac;(-a)(-b)=ab \text{ 和 } a\cdot 0=0.$$

韦伯更进一步注意到域中的积仅当它有一个因子是零时为零.然后他给出了域的几个例子,包括有理数域、有限域(只提到模一个素数的剩余类)和"形式域",即给定域 F 上的那些一个或多个变量的有理函数的域.但奇怪的是,域的这个抽象概念在韦伯的代数书中没有写入.显然,韦伯仍然认为更加具体的概念对学生来说更合适.

15.4　符号代数

在19世纪,英国的代数学以符号运算和它与数学事实的关系的新兴趣为特征(见补遗15.1和15.2).一个新的代数运动的领袖,更一般地说,沉迷于英国数学研究的变革的一个人是乔治·皮科克(1791—1858).皮科克在1830年他的《代数学》一书中介绍了他的代数新符号方法.这本书分别在1842和1845年扩充和修改.皮科克对改革的兴趣可以追溯到关于负数和虚数的意义的问题,它们已经被18世纪后期许多英国数学家提了出来.负数和虚数在18世纪(及以前)被自由运用,被认为是获得各种代数结果所必需的.但是数学家不能用各种物理意义之外的方式给出解释.就是因为这些概念缺少充分的基础导致特朗西斯·马塞雷斯(1731—1824)和威廉·弗兰德(1757—1841)所写的代数教材明确否认了它们的作用.但是,很明显,在方程的解的研究中给予负数和虚数以实际的意义这一步太激进而不能被广泛的接受.

补遗 15.1

剑桥的数学

到18世纪中叶,数学已经在剑桥研究制度中非常有影响力了,以至于剑桥最重要的考试——院考试即通常称为三脚凳的荣誉学位考试(tripos),因为原先选手们在回答提问时都坐过这种凳子而得名——主要是数学考试.毕竟,数学研究被看作是发展人的思维的,因而有助于为英国的绅士们获得在教会和政府中的领导权做准备,考试中的数学内容包括欧几里得和阿波罗尼奥斯在内的综合测试内容,其中包括代数学、三角学、流数以及牛顿的《原理》第一卷中的物理学基础.但是那些更认真的学生,即想成为 Wranglers(译注:剑桥大学荣誉学位考试的优胜者的称谓)——位于荣誉之首——的学生则自学更多前卫的数学,这些内容包括《原理》的其他内容以及在19世纪初增加的法国数学家(比如:拉格朗日,拉克鲁瓦和拉普拉斯)的著作.成为一位 Wranglers,基本上保证在剑桥有了学院奖学金,并因此而开始自己的事业,这对于任何生活未独立的学生来说都是重要的考虑.

在整个18世纪传统的数学教育模式就是综合几何法,甚至在牛顿的《原理》中也采用此方法.也正因如此,剑桥的学生很难理解在欧洲大陆上成功使用的分析方法.为了挽救这个状况,剑桥的几个本科毕业生1812年决定成立一个新的学会:分析学会.其目的是把欧洲的分析数学推广到英国,尤其是把这种数学推广使之成为剑桥的必修课.尽管该分析学会仅仅坚持了大约一年,它最初的许多成员,包括皮科克和查里斯·巴贝奇(1792—1871),对剑桥19世纪20年代中期转变到新的分析方向起到了重要的作用.他们工作的一个效果是完全改变了数学在剑桥的作用,即从全才教育向专才教育的转变,其目标是把新数学发展也作为知识进步的一部分.

补遗 15.2

1785 年剑桥荣誉学位考试

这份问卷是从一份备忘录中得出的，要求参加考试者书写清楚，并以“考生答案的清晰准确度评分而不是以答案的量来评分.”[41]

1. 证明有多少正立方体，它们叫什么，为什么没有其他更多的正立方体了.
2. 证明双曲线的渐进线总是在曲线的外面.
3. 假定从地球的某一高地向外抛出一物体，要用多大的速度抛出才能使之变为相对于地球的第二颗行星.
4. 证明在所有的圆锥截线上当力趋于焦点时它与距离平方成反比.
5. 设物体在不同的椭圆轨道上绕相同的力心作周期运动，其时间与平均距离平方的三次方成正比，证明力与距离的平方成反比.
6. 牛顿第三节和第七节的关系是什么?第三节是怎么运用到第七节的?
7. 把双二次方程 $x^4 + qx^2 + rx + s = 0$ 化成三次方程.
8. 求 $\dot{x} \times \sqrt{a^2 - x^2}$ 的流数.
9. 找一个数，从它中间取出它的平方将得到最大的差数.
10. 求圆周 $DBRS$ 的(任意) 一段弧 DB 的长度.

15.4.1　皮科克的《代数学》

皮科克自己认为他通过把代数区分成两类而拯救了负数和虚数. 他把代数分成了“算术代数”和“符号代数”，前者是通用算术，即一种通过利用字母而非数字本身来发展非负实数的基本代数性质的方式. 因而，在算术代数中人们可以写如下的式子：$a - (b - c) = a + c - b$，但只有在 $c < b$ 且 $b - c < a$ 的条件下才这样表示. 因而这个减法实际上是可以做的. 另一方面，在符号代数里，这些符号(字母) 不需要任何特别的解释. 这些字符的运算是从算术中类似的运算导出，但在符号代数中不必要限制它们的应用范围. 例如，上面的那个式子在符号代数中普遍成立.

皮科克对于什么是负数的问题的回答是：它仅仅是形式为 $-a$ 的一个符号，人们就像在算术里一样，对这些符号进行运算. 既然如果有 $a > b$ 且 $c > d$ 话，算术式子 $(a - b)(c - d) = ac - bc - ad + bd$ 成立，那么相同的规则在符号代数中也成立且不需要上述条件. 由此可得，置 $a = c = 0$ 时，$(-b)(-d) = bd$，置 $a = d = 0$ 时，$(-b) \times c = -bc$. 相似的，$\sqrt{-1}$ 就是一个满足与平方根在算术中遵守相同的规则的符号. 所以 $\sqrt{-1} \cdot \sqrt{-1} = -1$. 这些就是皮科克称之为等价形式的永久性原理的例子：“无论与什么代数形式等价，当这些符号一般为形式而不是确定的值的时候，将与这些符号为数值和形式时等价.”[42] 换句话说. 算术的任何可以用方程表示的法则，通过去掉所涉及到的符号的限制性，可确定符号代数的一个法则. 作为一个不含负数和虚数的例子，皮科克证明，既然

$$(1 + x)^m = 1 + mx + m(m - 1)\frac{x^2}{1 \cdot 2} + \cdots$$

当 m 为有理数且 $0 < x < 1$ 时在算术中成立，相同的方程在符号代数中也同样成立. 并且不论 x 和 m 为何都成立.

1830 年皮科克在他的《代数学》第一版中，把符号代数学定义为：“一门通过某种确定而又任意的方式处理符号和记号的组合的学科.”[43] 之后他开始把代数焦点从符号的意义上转移到了这些符号的运算规则上. 但是皮科克在 1830 年和 1845 年都没能利用他倡导的 “任意法则” 的组合. 事实上，他的符号代数里的所有的法则都是从相同运算的相应算术法则的“永久性原理” 导出的. 在

1845 年他写道:"我相信使组合法则随意且独立于运算的符号代数的特性没有显现可能是正确和有道理的."[44] 但是,不考虑他断言的"自由"为符号代数创造法则方面的失败,皮科克"关于这种代数的结果也许仅只能依约定而存在"[43] 的陈述标志了整个代数学科的一种新意义的开始.这种意义很快就被英国的其他数学家们所开拓.

人物小传	乔治·皮科克(1791—1858)(George Peacock)
	皮科克出生于林肯郡的一个小镇——登场,这里距牛顿的出生地只有几英里远.1809年,他进入剑桥大学三一学院学习,4年后以Wrangler第二的成绩毕业,并从三一学院的普通成员逐步成为学院讲师和导师,并于1837年成为天文学和几何学教授.正是由于他作为1817到1819年剑桥考试主管人的身份才使得把欧洲大陆的数学引入到考试中去.几年后,他参加了剑桥条理改写委员会,把宗教信仰测试是获得学位的必要条件之一这个要求去掉了.

15.4.2 德摩根和代数法则

奥古斯塔斯·德摩根(1806—1871)通过读皮科克的文章而受其影响,而且比他的前辈更清楚地认识到代数法则的创造不必用算术法则作为启发.德摩根相信代数系统可以不通过算术法则而通过任意一些符号和这些符号上的运算法则来创造.只要其后能解释这些法则就行了.在1849年,他给出了这种创造的一个例子:

> 假定有符号 $M,N,+$ 以及惟一的一种组合关系,即 $M+N$ 与 $N+M$ 相同.这里有一个符号演算:我们如何才能构造一个有意义的代数呢?如下是摘自众多方式中的几种:(1)M 和 N 可以是一些量,$+$ 就是这两个量的和的符号.(2)M,N 可以是数,那么"+"表示第一个数被第二个数乘.(3)M 和 N 可以是线,那"+"就表示由前一个线作水平边,后一个线作垂直边做一个矩形.(4)M,N 可以是人,那"+"就表示前一个人是后一个人的兄弟.(5)M,N 可以是国家,那么"+"表示后一个与前一个之间打过一场战争.[47]

尽管德摩根宣称对他的符号有创造代数公理的自由,甚至意识到这些符号可以不必是数值或者量,但是和皮科克一样,他也没有创造任何新的与算术中的数所遵守的规则不一样的规则体系.事实上,在1841年他建立了他相信的是这些"代数过程必不可少"的法则.这些法则包括替换法则(两个用 = 连接的式子可以用一个替换另一个),加法和乘法的逆运算法则(即 + 与 − 和 × 与 ÷ 在效果上是相反的),加法和乘法的交换律,加法对乘法的分配律以及减法对乘法的分配律和指数运算的规律

$$a^b a^c = a^{b+c}, (a^b)^c = a^{bc}.$$

提出他认为"既充分又不可去掉"这些规律之后,他指出:"最非凡的一点是前述的运算法则与相继的符号 $a+b$, ab 和 a^b 之间的联系比从也许认为是充分的算术解释中推导出的法则的联系要少得多."[49] 换句话说,没有必要让乘法的意义从加法而来,也没有意义让幂运算的意义从乘法的意义而来.然而,尽管一定可以用这些原理来推导所有的代数结果,但这样的代数并不比用每块拼图片的背面来做拼图游戏有更多的意思.德摩根相信真正的数学一定有真正的内涵.设计一个公理体系的结构的重要性远远不及解释说明它.德摩根认为公理所建立的逻辑框架之外,只有解释才能赋予数学体系以意义和重要性.

人物小传	奥古斯塔斯 · 德摩根(1806—1871)(Augustus De Morgan)
	德摩根出生于印度的一位英国军官家庭. 19 世纪 20 年代在剑桥大学学习,当时皮科克的改革已经初见成效,因而从一开始他就被教授了欧洲大陆的分析数学. 因为他在 1827 年仅以 Wrangler 第四的成绩毕业,部分是由于其他兴趣干扰了毕业考试必需的"死记硬背"的完成. 他觉得这个成绩太差以至于不敢再从事数学,因而开始在法院里谋职. 尽管如此,1828 年他还是被选为新创建的伦敦大学的数学主席. 他余生中大部分时间都是在这里度过的. 德摩根是一位尽职的老师,往往一学期要讲 4 门课,课程从基础的数学到变分学. 他把他的创造天才用到了设计新的教学方法上了,他不仅写了许多数学文章,而且还写了许多关于数学教学的文章和书.

15.4.3　哈密顿:复数和四元数

正是这位爱尔兰数学家和物理学家威廉 · 罗万 · 哈密顿(1805—1865) 最终创建了一个新的具有真正解释的代数系统. 这个系统没有满足由德摩根设立的公理. 就像皮科克和德摩根一样,哈密顿也想能证明负元和虚元在代数中的应用的合理性,他认为这些概念的基础不牢固. 正如在 1837 年他的重要论文《共轭函数论或代数偶;关于作为纯洁时代科学的代数方面的预备性的和基础性的随笔》中所写的:

> 当我们定义了下述规则(像平常一样) 就无需特别地怀疑和不相信负数和虚数的理论:一个较大的量可以被一个较小的量减,所得的差就比什么都没有还少,两个负数或者代表一些量可以相乘 …… 他们的积是一个正数 …… 无论一个数是正的还是负的它的平方总是一个正数. 那么虚数,也是可以找到、推出或者确定出来的,并在所有的正负数规则下运算. 它们好像也遵从上述的法则一样,虽然它们有负的平方. 因而它们不能看作正数、负数或空数了. 所以这个量既不表示比什么也没有大,也不表示比什么也没有小,甚至不表示等于什么也没有. [50]

为了使代数象几何一样建立在牢固的基础上,就需要为它创造一些直观的原则. 哈密顿认为可以从纯时间的直觉得到. 哈密顿的纯时间的思想源于他阅读伊曼纽尔 · 康德的《纯粹理性批判》中得到的启发:"它是我们内在观念的组成形成把感觉或者看法排列成同时或者顺次发生的内在感觉形式."[51] 用现代术语来说,哈密顿假定有一个"时刻"的集合 M,它被一个关系"$<$"所排序,对 M 中的任意元 $A,B,A>B,A=B,A<B$ 中必有一个成立,之后,他在时刻偶上定义等价关系. 如果序偶 (A,B) 和 (C,D) 满足下列的条件就说他们是等价的:"如果 B 与 A 完全相同,那么 D 必须与 C 也完全相同;如果 B 比 A 晚,那么 D 一定比 C 晚,而且晚的时间是一样的;如果 B 比 A 早,那么 D 一定比 C 早,而且早的时间是一样的."[52] 为了防止后面引起混乱,必须注意到哈密顿没有用现代的配对的符号来讨论这个等价. 他第一个使用具有提示性的符号 $B-A$ 来定义这个等价类的. 后来用了一个单个的符号 a,这个符号可以使人联想到有从 A 到 B 的**时间步长**的意义.

正是这个**时间步长**为哈密顿提供了构造负数的基础. 即,如果 α 描述了偶 $B-A$,那么 $\Theta\alpha$(哈密顿用来表示 $-\alpha$ 的符号) 就表示了偶 $A-B$. (哈密顿用拉丁文"相反"一词的第一个字符 Θ,创造了这个符号.) 假定步长 α 是单位,利用两个步长和的自然定义,哈密顿进一步构造了有理数集. 正数由 α 与它本身的乘积(连续和) 来定义,负数由 $\Theta\alpha$ 的倍数来定义,有理数通过比较步长 α 的两个整数倍来定义. 之后哈密顿阐明了这些(正的和负的) 倍数上算术运算的严格的规则. 比如:两个 α 的

负倍数的乘积一定是正的,因为这个乘积涉及把步长 a 的方向翻转两次.然而,由于对他上面所表示的负数甚为满意,他没有再考虑那个"比什么都没有还小"的量.哈密顿想进一步从有理数中构造出实数,这种尝试在现代观点看来不仅是一个失败,而且对德国所进行的分析数学算术化没有起到任何作用.但另一方面,他在这篇文章的最后一部分从实数把复数构造出来的方法被今天的许多教科书所采用.

在最后一部分中,哈密顿考察了时刻偶、时间步长以及数.因而两个时刻偶(A_1,B_1),(A_2,B_2)就决定了一个时间步长的偶$(a,b)=(B_1-A_1,B_2-A_2)$,由于两个步长偶$(\alpha a,\alpha b)$,(a,b)之间存在比例 α,从而哈密顿认为任何两个步长偶可表为数偶(α,β)的比.(以前的比 α 将由数偶$(\alpha,0)$代替.)很明显,这种数偶的加法以及减法将由下式定义:

$$(\alpha,\beta)\pm(\gamma,\delta)=(\alpha\pm\gamma,\beta\pm\delta).$$

假定乘法的分配律成立,哈密顿由此得乘法的一般规律可由下式给出:

$$(\alpha,\beta)(\gamma,\delta)=(\alpha\gamma-\beta\delta,\beta\gamma+\alpha\delta).$$

从而除法被定义为:

$$\frac{(\alpha,\beta)}{(\gamma,\delta)}=\left(\frac{\alpha\gamma+\beta\delta}{\gamma^2+\delta^2},\frac{\beta\gamma-\alpha\delta}{\gamma^2+\delta^2}\right).$$

正如哈密顿写道:"这些定义并不是随意选择的,尽管也可以取其他的定义,但没有任何定义能比这些定义更恰当."[53] 因为由这些定义可推出已知的那些复数运算法则.比如:$(0,1)(0,1)=(-1,0)$,如果把$(\alpha,0)$视为 α,因而$\sqrt{-1}$就可以定义为数偶$(0,1)$.复数 $\alpha+\beta\sqrt{-1}$ 可定义成数偶(α,β).上述的这些法则确定了复数运算的标准规则.哈密顿因此成功地绕过虚数而从实数中构造了复数,回答了复数究竟是什么这个问题.

哈密顿在他文章的最后评论说:"他希望能发展三元组以及时刻、步长、数所构成的集合的理论,其中包括二元数偶的理论."[54] 他已知道复数运算在二维空间已有一个几何解释.但是由于物理学的大多数事件发生在三维空间中,三元数的运算体系(即,一个代数)被证明是非常有用的.就像1841年他给德摩根的信中写到的:"如果我的代数观点是正确的话,以某种或者其他方式介绍三元甚至多元数是可能的.这样的数在某种意义上就满足符号方程 $a=(a_1,a_2,\cdots,a_n)$,在这里 a 是一个符号,表示一种复杂的思想,$a_1,a_2,\cdots,a_n$ 代表 n 个正的或负的实数."[55] 当然对哈密顿来说,他的困难不是三元数的加法(那是容易的),而是在乘法上.在他知道了数论的基本法则后,他希望三元数也能够满足乘法的结合律、交换律以及分配律、除了除数为零的情况外总可以做除法,并且结果惟一.同时他还想使模的性质也成立,即,如果$(a_1,a_2,a_3)(b_1,b_2,b_3)=(c_1,c_2,c_3)$成立,则有$(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)=c_1^2+c_2^2+c_3^2$.总之他希望各种运算在三维空间中有一个合理的解释.早在1830年他就开始为三元数的乘法寻找规则了,经过十三年的努力他终于解决了,但并不是以他所希望的方式,而是以本章开头所描述的那种方式.

哈密顿的解不是考虑三元数而是四元数(a,b,c,d),类似与复数标准记号就是 $a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$.基本的乘法规则为 $\mathrm{i}^2=\mathrm{j}^2=\mathrm{k}^2=\mathrm{ijk}=-1$,导出的规律为 $\mathrm{ij}=\mathrm{k},\mathrm{ji}=-\mathrm{k},\mathrm{jk}=\mathrm{i},\mathrm{kj}=-\mathrm{i},\mathrm{ki}=\mathrm{j}$ 和 $\mathrm{ik}=-\mathrm{j}$.当通过分配律把上述的规律扩展到所有的**四元数**上时,这个体系就具有了哈密顿所期望的除乘法交换律之外的所有的性质.在现代术语中,四元数集合形成了实数上的非交换的代数.哈密顿的这个体系是第一个不遵守皮科克和德摩根所建立的规律的量.他的创立为考虑那些不满足这些规律的体系清除了一个障碍,使皮科克所提倡的自由的创造成为了现实.

哈密顿本人非常重视他的发现,他用余下的时间写了好几卷关于四元数的理论.在这些著作中,通过考虑向量的商,他证明了在处理三维空间具有四个分量的物体时的合理性.向量 $\boldsymbol{w}$ 除向量

v 的商表示能把 **w** 变成 **v** 的量. 在二维空间中这个量由两个数值来表示:两个向量长度的比值和把 **w** 转到 **v** 需要的角度. 因而,可以很合理地认为二维空间中两个向量的商仍然是二维空间的向量. 但是在三维空间中旋转量本身依赖于三个数值,两个用来表示旋转轴的方向,一个表示旋转角度的大小,第四个量表示长度的比. 三维空间中两个向量的商决定了一个有四个分量的量,即一个四元数.

尽管,很少有物理学家用到四元数,但哈密顿的观念标志着向量概念在物理学理论中普遍应用的开始. 事实上,哈密顿特别提到把四元数 $Q = a + bi + cj + dk$ 写成两部分:实部 a 和虚部 $bi + cj + dk$ 是非常有好处的. 他把前一部分叫做**数量部分**,"因为它可以取 R 上的一系列数." 把后一部分称作**向量部分**,"因为它可以在三维空间中用几何的方法构造成直线或径矢."[56](径矢这一词早在 18 世纪早期就是一个数学名词了. 然而,是哈密顿首先在更广意义下用"向量"的.) 因此,哈密顿写成 $Q = S.Q + V.Q$,这里 $S.Q$ 是数量部分,$V.Q$ 是向量部分. 当我们考虑下面 α 和 β 的乘积:



图 15.5 爱尔兰邮票上的哈密顿.

$$(ai + bj + ck)(xi + yj + zk) = -(ax + by + cz) + (bz - cy)i + (cx - az)j + (ay - bx)k,$$

α、β 是数量部分为零的两个四元数,那么 $S.\alpha\beta$ 是向量 α 与 β 的现在所说的点积的负数,而 $V.\alpha\beta$ 是这两个向量的叉积.

人物小传	威廉 · 罗万 · 哈密顿(1805—1865)(William Rowan Hamilton)
	哈密顿出生于都柏林,但是在距东北部 30 英里的特里姆镇受的教育,由他的叔叔——一位传统学者指导. 由于哈密顿在幼年就显露了他的天赋,他的叔叔把他的多余的精力转移到语言学习上,到他刚刚 10 岁的时候,哈密顿不仅对拉丁语、希腊语和现代欧洲语言非常流利,而且对希伯来语、波斯语、阿拉伯语、梵文等其他语言也很流利了,要不是因为在都柏林汉语书很难找,他的叔叔也会让他精通汉语. 从很早的时候,哈密顿用他自己发明的计算方法开始学习数学,但是他对数学的兴趣是通过与一位美国计算奇人的联系而激发的,这位奇人在竞赛中一贯比哈密顿要好. 不久之后,他发现了欧几里得几何学和其他的许多数学领域. 到 1823 年他进入了都柏林的三一学院学习时,他已经打算研究在三一学院教授的大陆分析数学了. 这些课程是由于剑桥大学所倡导的改革而设的. 哈密顿很快就超过了设定的那些课程的范围,并掌握了综合工科学校的数学课程. 他的第一个原创工作是关于光的,而不是纯数学. 事实上,今天他在动力学上的名气比数学上的名气还大. 正是这一工作使他于 1827 年获得爱尔兰皇家天文学家的聘任. 当时,他甚至还没有拿到学位. 在这个职位上,他度过了自己的余生. 他在数学和物理学上的贡献使他在 1865 年成为新创建的英国国家科学院第一位外国合作伙伴.

15.4.4 四元数的向量

哈密顿关于四元数概念在物理学中的应用的继承者是苏格兰物理学家彼得 · 古日里 · 泰特(1831—1901)和詹姆斯 · 克拉克 · 麦克斯韦(1831—1879),他们在爱丁堡大学和剑桥大学中一直是朋友和同学. 泰特在 1867 年写了一篇提倡四元数方法在物理学中应用的文章《论四元数的基础理论》. 泰特的讨论本质包含了所有与现代意义下向量点积和叉积的规律相当的内容,尽管用成了四元数的记号. 特别的,他证明了 $S.\alpha\beta = -T\alpha T\beta\cos\theta$(这里 $T\alpha$ 表示 α 的长度,θ 是 α 和 β 之间的夹角),

和 $V.\alpha\beta = T\alpha T\beta\sin\theta \cdot \eta$(这里 η 是一个垂直 α 和 β 的单位向量).

麦克斯韦,在他的《论电和磁》中,同样提倡哈密顿的想法.但是,他的主要目的(正如他在他的序言中写到的)是为了避免引入复杂的笛卡尔坐标,立刻把注意力固定在空间中的某一点而不是它的三个坐标上,固定在它的大小和方向上而不是三个坐标上."[57] 因此,四元数和相关向量就要以一种比通常的坐标形式更加概念化的方式来表示物理量.在他的论述中,麦克斯韦把他的物理结论全部表示成两种形式,坐标和四元数两种形式.



图 15.6　为纪念四元数发现而发行的爱尔兰邮票.



图 15.7　圣马力诺邮票上的麦克斯韦.

但是,是耶鲁大学的吉布斯(1839—1903)和英国的亥维赛(1850—1925)在读了泰特和麦克斯韦的文章之后,分别独立地认识到整个四元数代数对讨论物理概念不是必不可少的.仅向量的两种乘积:点乘和叉乘就足够了.吉布斯在 1881 年和 1884 年非公开地出版了他向量分析的结果,并在耶鲁大学讲了许多年这个课.亥维赛于 1882 年和 1883 年在他的关于电的论文中第一次发表了该方法.但是,我们所用来表示点乘和叉乘的符号 $A \cdot B$ 和 $A \times B$ 是属于前者的.随着 1901 年吉布斯的从演讲中获得的《向量分析》的正式出版,物理界更加清楚地知道是向量而不是四元数为描述物理概念提供了必需的语言.尽管四元数在数学上还是很重要的,但它在物理学中的应用已悄悄地消失了.

15.4.5　布尔和逻辑学

由皮科克和德摩根所提倡的代数的自由性由英国自学的逻辑学家乔治·布尔(1815 – 1864)用另一种方式进行了探索.1847 年,布尔出版了一本小书《逻辑的数学分析》,七年以后,把它扩展成了一本《思考规律的研究》,这本书使得逻辑学的研究最终脱离了形而上学而进入了数学.逻辑学从亚里士多德时代就被认为是形而上学.布尔的《思考规律》的目的是"研究人脑的用来完成推理的那些运算的基本规律,并给出符号语言的运算表达式,在这基础上建立逻辑学并构造它的方法."[58]

由于代数学是通过符号来研究的,布尔把在逻辑学中将被分析的那些基本符号放在了他的第一个命题中:

命题 1　作为推理的一种手段,这个语言所有的运算可能被由下述符号所组成的系统来指导:

1. 文字符号,如 x,y 等,代表我们设想的主题物.

2. 运算符,如 +、-、×,代表人脑的运算,通过这些运算,事物的概念被组合和融合,从而形成涉及同样元素的新概念.

3. "相同的"符号,=.

这些逻辑符号遵从确定的法则,与代数学中相应符号的运算法则相比部分相同,部分不同.[59]

布尔下一步定义他的语言符号的法则,一个字母可以代表一个物体的类或集合.因此 x 可以代表"男人"这一类,而 y 代表"好东西"这一集合.他们的组合 xy 将代表 x 和 y 都适用的一个类.即:

"好男人"类.显然,布尔的乘法满足交换律:$xy = yx$.在其他的乘法规则中,布尔得到 $x^2 = x$,因为对 x 和 x 都适用的类就是仅对 x 适用的类.如果 x 所代表的类包含在 y 所代表的类中,则 $xy = x$.

对布尔来说,加法写成 $x + y$,代表两个类 x 和 y 的连接,减法写成 $x - y$,代表那么由在 x 类中且不在 y 类中的元组成的类,交换律和分配律也成立,$z(x \pm y) = zx \pm zy$.用 0 表示空类,1 代表全体组成的类.布尔又非常类似地得到了熟悉的法则 $0y = 0$ 和 $1y = y$.同时还有不太熟悉的规则 $x(1 - x) = 0$,该式"断定任何满足某性质同时又不满足该性质是不可能的."[60]

由于上述法则限制在 0 和 1 上的数的法则一致,布尔决定他的逻辑代数将处理那些只能取这些值的变量.特别的,他考虑一个或几个逻辑变元的函数 $f(x), f(x,y), \cdots$.在这里,变量只能取值 0 或 1.例如,他证明任何这样的函数都可展成 $f(x) = f(1)x + f(0)(1 - x)$ 的形式.或者设 $\bar{x} = 1 - x$,写成 $f(x) = f(1)x + f(0)\bar{x}$,类似的,$f(x,y) = f(1,1)xy + f(1,0)x\bar{y} + f(0,1)\bar{x}y + f(0,0)\overline{x}\overline{y}$.例如,函数 $1 - x + xy$ 可以写成 $1xy + 0x\bar{y} + 1\bar{x}y + 1\overline{x}\overline{y} = xy + \bar{x}y + \overline{x}\overline{y}$.

布尔进一步证明如果 V 是某一个函数,我们可以根据以上规则把 V 展开而解释方程 $V = 0$,并让每个系数不为 0 的部分等于 0.布尔举了犹太法则中的"干净的野兽"的定义为例子,干净的野兽总是蹄子分开且反刍的动物.用 x、y、z 分别代表干净的、蹄子分开、和反刍的动物,那么干净的动物定义成方程 $x = yz$ 或者 $V = x - yz = 0$,展开 $x - yz$,布尔得到

$$V = 0xyz + xy\bar{z} + x\bar{y}z + x\,\overline{y}\overline{z} - \bar{x}yz + 0\bar{x}y\bar{z} + 0\,\overline{x}\overline{y}z + 0\,\overline{x}\overline{y}\overline{z}.$$

令每个系数不为零的项都等于零,

$$xy\bar{z} = 0 \quad x\bar{y}z = 0 \quad x\,\overline{y}\overline{z} = 0 \quad \bar{x}yz = 0.$$

这些等式的意义就是某些类型的动物不存在.比如,第一个式子是说不存在干净的、蹄子分开但不反刍的动物.

布尔所发展的类的代数就是我们今天所知道的布尔代数,他好像在发表之后的很多年都没有再发展.它现已成为研究电路设计代数研究中的中心,在现代计算器和计算机的逻辑代数得到了发展.布尔如果看到他对思维规律的运算几乎以一个世纪前他预言的方式发展将会很高兴.

15.5　矩阵和线性方程组

矩阵的思想很早就已经有了,至少可以追溯到汉代中国学者在解线性方程时的应用,在 18 世纪或更早些时候,数的方阵的行列式已被计算和使用了,通常是在解线性方程组时使用.尽管在当时阵列本身并没有单独引起注意.19 世纪的其他工作导致阵列更加形式的计算,并在 19 世纪中叶导致了矩阵概念的定义以及矩阵代数的发展.除了这些形式化的工作,还有矩阵论发展中深刻的一面,即从高斯二次型的研究中发展出来的成果,并最终引起了相似、对角化和标准型的矩阵分类.

15.5.1　矩阵的基本思想

高斯在他的二次型理论中讨论到了可以把一个形式转化成另一个形式的线性变换的思想,如果 $F = ax^2 + 2bxy + cy^2$,那么变换 $\begin{cases} x = \alpha x' + \beta y' \\ y = \gamma x' + \delta y' \end{cases}$ 把 F 变成一个新的形式 F',它的系数依赖于 F 的系数和变换本身,高斯特别指出如果 F' 通过另外一个变换 $\begin{cases} x' = \varepsilon x'' + \xi y'' \\ y' = \eta x'' + \theta y'' \end{cases}$ 变成 F'',那么这两个变换的复合就是一个把 F 变成 F'' 的新的变换:

$$\begin{cases} x = (\alpha\epsilon + \beta\eta)x'' + (\alpha\xi + \beta\theta)y'', \\ y = (\gamma\epsilon + \delta\eta)x'' + (\gamma\eta + \delta\theta)y''. \end{cases}$$

这个新变换的系数矩阵是原来的两个变换的系数矩阵的乘积，高斯在他研究三元二次型 $Ax^2 + 2Bxy + Cy^2 + 2Dxz + 2Eyz + Fz^2$ 时，演算了一个相似的计算过程，这实际上给出了 3×3 矩阵相乘的法则. 尽管他把变换的系数写成矩形阵列的形式，甚至用单个的字母 S 指代一个特殊的变换，高斯并没有明确指出这种复合的思想就是乘法.

在 1815 年，柯西发表了一篇关于行列式理论的基础性文章. 在这篇文章中他不仅用这个名字代替了几个旧的术语，而且用缩写的记号$(a_{1,n})$代表他称之为"对称组"的矩阵：

$$\begin{matrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{matrix}$$

与它有一个相关的行列式. 尽管关于行列式计算的许多基本结论很早就已经知道了，是柯西第一次在他的论文中给出了它们的完整论述，包括一个给定的矩阵伴随余子式的思想以及通过展开任何行或者列来计算行列式的步骤，继而在高斯之后明确地认识到复合两个组$(a_{1,n})$和$(b_{1,n})$可以得到一个新的组$(m_{1,n})$的思想. 后者是通过熟知的乘法律定义的：

$$m_{i,j} = \sum_{k=1}^{n} a_{i,k}b_{k,j}.$$

之后，他证明了新组的行列式是原来两个组的行列式的乘积.

高斯的一位于 1843 年到爱尔兰拜访过哈密顿的学生费尔迪南 · 奇特候德 · 爱森斯坦(1823—1852)用明确的符号 $S \times T$ 来表示两个变换S和T的复合. 这些内容写在1844年他的一篇讨论三次型的论文中，也许是从高斯的行列式乘法定理中得到的启发. 关于这个记号，爱森斯坦写道："顺便地，在它的基础上可以建立一个算法，其中包括把乘、除法以及乘幂的一般运算规则应用到两个线性方程组的符号方程上. 正确的符号方程总是可以得到，它思考的中心问题是因子的顺序，即，方程组复合的顺序往往不可以改变."[61] 推测一下这个乘法不可交换的代数体系是否与爱森斯坦在 1843 年和哈密顿的讨论有关将非常有趣，但可能永远也得不到答案.

15.5.2 矩阵运算

爱森斯坦没能充分发展他的变换的代数的思想是因为他在 39 岁时就去世了，这方面的进展是由英国的亚瑟 · 凯莱和詹姆斯 · 约瑟夫 · 西尔维斯特(1814—1897)在 19 世纪 90 年代做出的.

在 1850 年西尔维斯特创造了**矩阵**一词来表示"一项由 m 行 n 列元素组成的矩形排列，"因为由那个排列，"我们能形成各种行列式组."(矩阵的英语原意是指可以引起其他事物的源头.) 他当时并没有用这个术语，是他的朋友凯莱在 1855 年和 1858 年的论文中用到了该词，在 1855 年的那篇论文中凯莱注意到线性方程组中使用矩阵是非常方便的. 因而他用：

$$(\xi, \eta, \zeta, \cdots) = \begin{pmatrix} \alpha, & \beta, & \gamma, & \cdots \\ \alpha', & \beta', & \gamma', & \cdots \\ \alpha'', & \beta'', & \gamma'', & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix} (x, y, z, \cdots)$$

来表示方程组

人物小传	詹姆斯·约瑟夫·西尔维斯特(1814—1897)(James Joseph Sylvester)
	西尔维斯特出生在伦敦的一个犹太人家庭,并且在剑桥学习过几年,但是由于宗教的原因他没能在这里获得学位.因此,他的学位是于 1814 年从都柏林的三一学院获得的,并在弗吉利亚大学获得了一个教授职位.由于他对奴隶制的厌恶,并与一个他认为不尊敬他的学生发生争执致使他于 1843 年辞职.回到伦敦之后先做了 10 年的律师,而后在伍尔里奇的皇家军事学院作了 15 年的数学教授,最后于 1871 年到巴尔迪摩新成立的约翰斯·霍普金斯大学任数学系主任,在这里,他创建了《美国数学杂志》,并帮助发展了美国对研究生数学教育的传统.

$$\begin{aligned}\xi &= \alpha x + \beta y + \gamma z + \cdots,\\ \eta &= \alpha' x + \beta' y + \gamma' z + \cdots,\\ \zeta &= \alpha'' x + \beta'' y + \gamma'' z + \cdots,\\ \cdots &= \cdots + \cdots + \cdots + \cdots.\end{aligned}$$

继而,他把方程组的解用矩阵的逆来表示:

$$(x, y, z, \cdots) = \begin{pmatrix} \alpha, & \beta, & \gamma, & \cdots \\ \alpha', & \beta', & \gamma', & \cdots \\ \alpha'', & \beta'', & \gamma'', & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{pmatrix}^{-1} (\xi, \eta, \zeta, \cdots).$$

这种表示法是把矩阵方程与只含一个变量的简单方程类比而得来的.但是,凯莱在知道了克莱姆的法则之后,把逆矩阵的元素用含有适当的行列式的分数来表示.在 1858 年,凯莱用了单个的字母表示矩阵,并给出了矩阵相乘、相加以及相减的规则:

> 我们很容易发现,同阶的矩阵和单个的量非常相似,它们可以被加、减或复合到一起.并且其加法和一般代数量的加法非常相似.考察它们之间的复合,发现矩阵是不可交换的.尽管如此,构造一个(正的或负的,整数的或小数的)矩阵的乘幂还是可能的,因而就有矩阵的有理函数与整函数,或者更一般的任何代数函数.[63]

之后,凯莱继续挖掘他的思想,不断地利用一般的代数运算和矩阵运算之间的相似性,并仔细留意这种相似性不成立的情况.因而利用了 3×3 矩阵的逆的公式,他写道:"当行列式变成 0 的时候,逆矩阵的概念就没有了,这种矩阵称为不定的,0 矩阵是不定的,只有当两个矩阵中的一个或两个都是不定时,它们的乘积才可能是 0."[64]

大概是由于凯莱用了单个的符号来表示矩阵才推出了所谓的凯莱–哈密顿定理.对 2×2 矩阵 $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,凯莱把他的结果明确地写成

$$\det\begin{pmatrix} a - M & b \\ c & d - M \end{pmatrix} = 0.$$

凯莱首先是在 1857 年 11 月份一封给西尔维斯特的信中表述这个"非凡"的定理的.他只简单地给出 $M^2 - (a + d)M^1 + (ad - bc)M^0 = 0$($M^0$ 指单位矩阵)就证明了上述结论.其叙述在本质上是现代形式的一般结论,即 M 满足 λ 的**特征方程**,$\det(M - \lambda I) = 0$.凯莱注意到他已验证了 3×3 矩阵的情形,进一步他写道:"我认为没有必要去验证一般的任意阶的矩阵满足这个定理了."[65] 由乔治·弗罗贝尼乌斯(1849—1917)利用凯莱新符号在 20 年之后给出完整证明.

凯莱提出凯莱 - 哈密顿定理的动机就是想证明:"任何矩阵都会满足一个与它同阶的代数方程," 并由此而得到"矩阵的任何有理整函数可以表示成阶不超过矩阵的阶的一个有理整函数."[66] 凯莱甚至继续证明这个结论适用于非有理函数.特别地,他给出了如何计算 $L = \sqrt{M}$,这里的 M 是上述的 2×2 的矩阵.结论是形如

$$\begin{pmatrix} \dfrac{a+Y}{X} & \dfrac{b}{X} \\ \dfrac{c}{X} & \dfrac{d+Y}{X} \end{pmatrix}$$

的矩阵.其中 $X = \sqrt{a + d + 2\sqrt{ad - bc}}$, $Y = \sqrt{ad - bc}$.但是凯莱没能给出这个结论的使用范围.对仅依赖于符号运算的类似的讨论,由于没有考虑到运算不成立的特殊情况,凯莱得出所有的与 M 可交换的矩阵 L 的错误特征描述,也正是同样的问题使约当于 10 年之后通过今天被称作的约当标准型而把矩阵进行了基本分类.

15.5.3 特征值与特征向量

约当的分类不是基于矩阵的形式运算,而是谱理论,即围绕着特征值的概念的结果.在现代术语中,矩阵的**特征值**是矩阵方程 $AX = \lambda X$ 的一个解 λ,这里 A 是一个 $n \times n$ 矩阵,X 是一个 $n \times 1$ 矩阵,或者 $XA = \lambda X$,这里 A 是一个 $n \times n$ 矩阵,X 是 $1 \times n$ 矩阵,特征向量是与 λ 对应的那个满足同一方程的向量 X.自始至终,这些概念都是独立于矩阵理论自身,它们从不同思想的研究发展而来,并最终包含在理论之中.因此,在 18 世纪常系数的线性微分方程组解的问题最早引起特征值的问题.在达朗贝尔的从 1743 年到 1758 年的著作中,由于考察了承载有限质量的运动而受启发,考察了下述方程组(为简单起见,把有限元的个数定义为 3 个):

$$\frac{d^2 y_i}{dt^2} + \sum_{k=1}^{3} a_{ik} y_k = 0 \quad i = 1,2,3.$$

为了解这个方程组,对三个方程分别乘上一个常量 v_i,而后加在一起得:

$$\sum_{i=1}^{3} v_i \frac{d^2 y_i}{dt^2} + \sum_{i,k=1}^{3} v_i a_{ik} y_k = 0.$$

如果所选的 v_i 使得

$$\sum_{i=1}^{3} v_i a_{ik} + \lambda v_k = 0 \quad k = 1,2,3,$$

即,如果 (v_1, v_2, v_3) 是矩阵 $A = (a_{ik})$ 相应于 λ 的特征向量,那么变换 $u = v_1 y_1 + v_2 y_2 + v_3 y_3$ 就把方程化简成单个的微分方程

$$\frac{d^2 u}{dt^2} + \lambda u = 0.$$

有了欧拉对微分方程所作的工作,这个方程就很容易解出来了,得到三个 y_i.通过对这三个方程的研究得到 λ 是由一个有三个根的三次方程决定,达朗贝尔意识到要使这个解有物理意义,当 $t \to +\infty$ 时,它们必须有界,继而,只有当三个 λ 是相异的正实数时才有意义.

柯西首先解决了特殊情况下通过矩阵本身的属性来确定特征值的属性的问题.总的说来,他没有受到达朗贝尔的关于微分方程的研究的影响,而是通过二次曲面的研究受到启发.这门课作为解析几何的重要部分,柯西从 1815 年开始就在综合工科学校讲授了,一个中心在原点的二次曲面由一个方程 $f(x, y, z) = k$ 给出,这里 f 是一个二次型,为弄清这种曲面,与欧拉一样,柯西需要找到

一个坐标变换使 f 变成一个只含平方项的形式.在几何中,这个问题就是在三维空间中找到一组正交的坐标轴,在这个坐标中表示这个曲面.柯西后来把这个问题推广到了有 n 个变量的二次型中,其系数可写成一个对称矩阵.例如,二元二次型 $ax^2+2bxy+cy^2$ 定义了 2×2 的对称矩阵 $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$.柯西后来的目的是要找到一个变量的线性变换,使矩阵在这个线性变换的作用下是对角化的.这个目标在 1829 年的一篇文章中达到了,由于一般情形下的细节有些复杂,而且柯西的证明实质在 2 个变量的情形下已很明显,因而我们在此只看到两个变量的情形.

为了找到一个把 $f(x,y)=ax^2+2bxy+cy^2$ 转化成平方和的形式的线性变换,只须找到 $f(x,y)$ 在 $x^2+y^2=1$ 的条件下的最大值和最小值,使 f 取极值的单位圆上的点同时也在由方程 $f(x,y)=k$ 所确定的某椭圆(或双曲线)的某个轴的顶点上.如果把原点到这个点的直线取作一个轴,与之垂直的直线取作另一个轴的话,在这个坐标系下,这个方程就只包含平方项了.由拉格朗日的乘数法,极值只有当 $\frac{f_x}{2x}=\frac{f_y}{2y}$ 成立时才取得,让两边等于 λ,得到两个方程

$$\frac{ax+by}{x}=\lambda,\frac{bx+cy}{y}=\lambda,$$

又可以写成方程组:$\begin{cases}(a-\lambda)x+by=0,\\ bx+(c-\lambda)y=0.\end{cases}$ 柯西知道只有当行列式等于 0 的时候,这个方程组才有非平凡解,即 $(a-\lambda)(c-\lambda)-b^2=0$,在矩阵的术语中,这个等式就是特征方程 $\det(A-\lambda I)=0$, 30 年以后,凯莱就是讨论的这个方程.

为了弄清特征方程的根是如何把一个矩阵对角化的,设 λ_1,λ_2 是方程的两个根,(x_1,y_1), (x_2,y_2) 是相应的解,因而,

$$(a-\lambda_1)x_1+by_1=0,\quad (a-\lambda_2)x_2+by_2=0.$$

用 x_2 乘前式减去后式乘以 x_1 的积,得方程

$$(\lambda_2-\lambda_1)x_1x_2+b(y_1x_2-x_1y_2)=0.$$

类似地,从包含 $(c-\lambda_i)$ 的方程开始,可以得到方程

$$b(y_2x_1-y_1x_2)+(\lambda_1-\lambda_2)y_1y_2=0.$$

这两个方程相加得:$(\lambda_2-\lambda_1)(x_1x_2+y_1y_2)=0$.因而,如果 $\lambda_2\neq\lambda_1$(在我们考虑的情形下无疑是成立的,除非原来的二次型已被对角化了),那么,$x_1x_2+y_1y_2=0$.因为 (x_1,y_1),(x_2,y_2) 可确定到差一个常数倍,可以要求 ${x_1}^2+{y_1}^2=1,{x_2}^2+{y_2}^2=1$.用现代术语来说,线性变换:

$$\begin{cases}x=x_1u+x_2v,\\ y=y_1u+y_2v\end{cases}$$

是正交的.很容易算出这个变换下得到的二次型正是所需要的 $\lambda_1u^2+\lambda_2v^2$. λ_1 和 λ_2 由假定可得是实的,如若相反则它们互为共轭,在这种情况下,x_1 是 x_2 的共轭,y_1 是 y_2 的共轭,$x_1x_2+y_1y_2$ 就不能为 0 了.因而,柯西证明了所有对角矩阵的特征向量都是实的,至少在不等的情况下是这样的,并且矩阵可以通过正交变换而对角化.

15.5.4　标准形

柯西论文中的讨论为处理各种类型的矩阵的特征值和标准型提供了一种广泛的理论.然而,几乎在整个 19 世纪中期这些理论都是用二次型的形式来表示的,而不是矩阵.二次型可以生成对称

矩阵.更加一般的是双线性型,形如$\sum_{i,j=1}^{n} a_{ij}x_i y_j$ 的含有 $2n$ 个变量的函数,生成一般的方阵.

关于型的理论的有影响的部分是由约当在他的《论变换》一文中解决的.约当没有通过双线性型而是直接用线性变换本身的研究,得到分类问题的解决.他曾经仔细研究过伽罗瓦的关于代数方程的解工作,尤其是关于素数幂次代数方程的解的理论.这些解与根的线性变换有关,系数可以看作是 p 阶的有限域的元素.这些关于根 $x_1, x_2, \cdots, x_n$ 的变换可表为矩阵A 的形式.换句话说,如果用 X 代表根x_i 的$n \times 1$矩阵,那么变换就可以写成 $X' \equiv AX(\mathrm{mod}\ p)$.约当的目的是想找到他所谓的"指标的变换",使之变换成尽可能简单的形式.写成矩阵的符号就是他想找到一个 $n \times n$ 的可逆矩阵 P,使得 $PA \equiv DP$(这里 D 是尽可能简单的矩阵).因此,若 $Y \equiv PX$,那么 $PAP^{-1}Y \equiv PAX \equiv DPX \equiv DY$,并且这个关于 Y 的变换是"简单的".利用 A 的特征多项式,约当指出如果 $\det(A - \lambda I) \equiv 0$的所有的根不同,那么 D 就可以取对角阵,其中对角线的元素是其特征值,另一方面,如果有重根,约当证明了可以找到一个变换使 D 分块成形如

$$\begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_m \end{pmatrix}$$

的矩阵,其中每个块 D_i 都是形如

$$\begin{pmatrix} \lambda_i & 0 & 0 & \cdots & 0 & 0 \\ \lambda_i & \lambda_i & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & \lambda_i \end{pmatrix}$$

的矩阵,$\lambda_i \not\equiv 0(\mathrm{mod}\ p)$ 是特征方程的根.这种标准型就是我们今天的约当标准型,对角线的元素用 1 来代换.1871 年当他发表这个结论时就意识到该方法可以用到线性微分方程组中去,这里的系数不是 p 域中的元素而是实数或复数.由此,继达朗贝尔 100 多年后,约当回到了与矩阵特征值相关的整个思想体系的源头.

约当并没有用凯莱的单个字母来代表线性变换.弗罗贝尼乌斯在 1878 年把前人的思想综合成一本完整的矩阵理论的著作.弗罗贝尼乌斯指出了各种不同类型的矩阵的关系.例如,他定义了两个矩阵 A 和 B 是**相似的**,如果有一个可逆矩阵 P 使得$B = P^{-1}AP$.如果存在 P 使得$B = P^{\mathrm{T}}AP$ 成立,就说 B 和A **相合**,这里 P^{T}是 P 的转置.他证明了,如果两个对称矩阵相似,那么变换矩阵可以取为**正交**的(一个矩阵的逆等于它的转置).而后,弗罗贝尼乌斯证明它们的特征值是模为 1 的复数,同时还证明了一些其他的结论,在论文的最后,弗罗贝尼乌斯说明了符号矩阵理论和四元数理论之间的关系.他定义了 4 个 2×2 的矩阵,这四个矩阵生成的代数刚好是四元数 1,i,j 和 k 生成的代数.

15.5.5 方程组的解

弗罗贝尼乌斯还曾尝试着弄清楚线性方程组解集的特征;欧拉在多年前考虑过它的一个特殊情形,当时他发现了一个没有惟一解的方程组,对此他感到困惑.当时他没有意识到是由于这个方程组的系数矩阵的行列式等于了0而出现了上述的情形.在19世纪中期,数学家们面临着各种各样的问题.他们不仅想弄清楚什么条件下 m 个 n 元线性方程组有解,而且想弄明白解集的大小.由行列式的经验可得,如果从这个方程组里取一个含 k 个方程的子组,它的 $k \times k$ 的系数矩阵的行列式

非 0,那么该子组就可解,尽管它不一定是惟一的.因此,存在原方程组行列式可以决定解集性质、以及该方程组是否可解的条件.在方程个数比变量个数多的情况下一般无解,因为它是超定的.因此我们只关心 $n \geqslant m$ 的方程组,用凯莱的符号记为:$AX = B$,A 是 $m \times n$ 的矩阵.

亨利 J.S.史密斯(1826—1883)是此后第一个在该理论中做出贡献的科学家之一,他是牛津大学的一位几何学教授.在1861年的论文中他发展了齐次线性方程组 $AX = 0$ 的不定指标和方程组的独立解的完全集两个基本概念.这两个概念都只处理 $n > m$ 并且 A 有一个 m 阶非 0 行列式的情形.$n - m$ 表示变量个数超过"独立方程的个数"的数.之后,史密斯证明存在一个含有 $n - m$ 个元的解集$(x_{i1}, x_{i2}, \cdots, x_{in})$ $i = 1, 2, \cdots, n - m$,其他的任何解都是这些解的线性组合,且阶为 m 的矩阵(x_{ij})的行列式的值不全为零.事实上,他找到了大量实际确定独立解的完全集的方法,并进一步指出要解决非齐次线性方程组 $AX = B$,此时 B 不等于零,只需找到一个特解 X^*,任何解就可以表示成 $X^* + X$ 的形式,X 是对应的齐次线性方程组的解.

尽管史密斯用到了"实际的独立方程",但他并没有考虑这个数比实际方程的个数小的情形.即,矩阵的非零子式的最大阶数小于 m 的情形.这个问题是由查理斯·L·道吉森(1832—1898)解决的,就是我们今天所说的L.卡罗尔.他在1867年发表的《行列式初论》一书中不仅讨论了 $AX = B$ 的 $m \times n$ 矩阵 A,还讨论了该方程组的 $m \times (n+1)$ 阶的**增广矩阵**$(A|B)$,它蕴含了该方程组是不是相容的.进而,他提出并证明了一个确定任意方程组解集性质的一般定理:

道吉森定理　设有 m 个 n 元方程组成的方程组($n \geqslant m$),存在 r 个方程使得 r 阶的增广矩阵的行列式非零,且当从余下的方程中再任取一个所组成的 $r+1$ 个方程的增广矩阵的每个 $r+1$ 阶行列式总为零,那么,该方程是相容的,如果 r 个方程的任一 r 阶非零行列式已选定,那么剩下的系数不在这个行列式中的 $n - r$ 个变量就可取任意值,对于这样的任意一组值,其他变量有惟一的值与它对应,而且依赖于这 r 个方程.[67]

道吉森的证明极具构造性,并随后给出了几个例子.下面考察含 5 个变量 4 个方程的方程组:

$$\begin{aligned}
u + v - 2x + y - z &= 6,\\
2u + 2v - 4x - y + z &= 9,\\
u + v - 2x \qquad\qquad &= 5,\\
u - v + x + y - 2z &= 0.
\end{aligned}$$

对于前两个方程有一个二阶的非零行列式,但对于前三个方程就没有 3 阶的非零行列式了,第 1、2、4 方程含有 3 阶的非零行列式,因此,道吉森断言这个方程组是相容的.由于它含有 5 个变量,有 $5 - 3 = 2$ 个自由变量,方程 3 是依赖于方程 1 和 2 的.

该定理已隐含了秩的思想,但他并没有抽象出来,这项工作是 1879 年由弗罗贝尼乌斯从其先辈工作中得来的,他写道:"如果一个行列式的所有 $r + 1$ 阶子式为 0,但至少有一个 r 阶子式不为零,那么就称 r 为行列式的**秩**."[68] 早在几年前,他明确了史密斯的"真正独立的方程"的意义,并把这种看法定义为方程和表示方程组的 n 元组的**线性无关性**;因而,在齐次方程组中,当 $c_1x_{1j} + c_2x_{2j} + \cdots + c_kx_{kj}$ 对于所有的 $j = 1, 2, \cdots, n$ 都不等于零时,解$(x_{11}, x_{12}, \cdots, x_{1n})$,$(x_{21}, x_{22}, \cdots, x_{2n})$,$\cdots$,$(x_{k1}, x_{k2}, \cdots, x_{kn})$是线性无关的($c_i$ 不全为零).为了定义方程的无关性,弗罗贝尼乌斯建立了一个对偶关系,对于给定的齐次线性方程组,有一个新的方程组,这个方程组的系数组成原方程的解的一个基.因此,n 元组和方程是从两个不同的视点看的同一事物.而后他证明如果一个由 m 个 n 元方程组成的方程组的秩为 r,就可以找到含 $n - r$ 个无关的解集,把系数和解集的坐标的位置转换一下,他发现了伴随方程组的秩为 $n - r$,并证明了这个方程组有一个与原方程组同解的伴随方程组.

例如:对同一其次方程组

$$\begin{aligned} u + v - 2x + y - z &= 0, \\ 2u + 2v - 4x - y + z &= 0, \\ u + v - 2x &= 0, \\ u - v + x + y - 2z &= 0, \end{aligned}$$

其秩为 3,因而有两个无关的解,这个解集的基可以取为(1,3,2,0,0) 和(1, -1,0,2,2,).伴随组就是

$$\begin{aligned} u + 3v + 2x &= 0, \\ u - v + 2y + 2z &= 0. \end{aligned}$$

这个方程组的秩为2,它的解集的基是(1,1, -2,0,0),(3, -1,0, -2,0) 和(3, -1,0,0, -2).可以直接地看出伴随这些解的这个方程组,即

$$\begin{aligned} u + v - 2x &= 0, \\ 3u - v - 2y &= 0, \\ 3u - v - 2z &= 0 \end{aligned}$$

和原方程同解.

尽管弗罗贝尼乌斯已经完成了方程组的解和各种特殊类型的矩阵的性质的研究,但是直到20世纪初期才出现了用矩阵术语来组织材料的教科书,并且直到40年代才认识到矩阵和向量空间的线性变换之间的关系,为了做到这点,必须把向量空间的抽象思想明确化.这方面的发展产生于某些几何思想,因而我们把它放到第 17 章中讲述.

习　题

高斯数论的问题

1. 证明如果 p 是满足 $0 < a < p$ 的素数,那么使 $a^m \equiv 1 (\bmod p)$ 成立的最小指数 m 一定是 $p-1$ 的因子.
2. 当 $p = 7$ 时,对每个 $0 < a < p$ 的整数 a,求使 $a^m \equiv 1(\bmod p)$ 成立的最小指数 m.并证明练习 1 中的定理对所有的 a 都成立.
3. 求 $p = 13$ 的原根,即,找 a 使 $p-1$ 是使 $a^{p-1} \equiv 1(\bmod p)$ 成立的最小指数.
4. 证明 -2 是形如 $8n+3$ 的素数的二次剩余,2 不是.
5. 按下面这几条完成高斯确定 453 是 1236 的二次剩余的断言的证明.证明:
   (a) 如果 $x^2 \equiv 453(\bmod 4)$, $x^2 \equiv 453(\bmod 3)$ 和 $x^2 \equiv 453(\bmod 103)$ 都可解,那么 $x^2 \equiv 453(\bmod 4 \times 3 \times 103)$ 也可解.
   (b)453 是模 4 和 3 的二次剩余.
   (c) $\left(\frac{453}{103}\right) = \left(\frac{41}{103}\right)$.
   (d) $\left(\frac{5}{41}\right) = 1$.
6. 证明高斯整数 $a + b\mathrm{i}$ 是素的,当且仅当模 $a^2 + b^2$ 是通常的素数.
7. 证明任何整除高斯数的乘积 $abc\cdots$ 的高斯素数 p 一定等于其中的一个素数或其中一个与单位元的乘积(提示:两边取模).
8. 分解 $3 + 5\mathrm{i}$ 为高斯素数的乘积.

代数数论的问题

9. 证明素的复整数是不可约的.

10. 在形如 $a+b\sqrt{-17}$ 的整数环中证明刘维尔分解式 $169=13\times13=(4+3\sqrt{-17})(4-3\sqrt{-17})$ 说明了惟一因子分解在这个环内不成立(提示:用模证明这四个因子是不可约的).
11. 证明高斯整数构成了一个欧几里得环,即证明对于给定的两个高斯整数 z 和 m,存在另两个数 q 和 r,使得 $z=qm+r$ 成立,且 $N(r)<N(m)$.
12. 证明形如 $a+b\sqrt{-2}$ 的复整数域是欧几里得的.首先具体确定这个环中类比的欧几里得算法.
13. 证明在形如 $a+b\sqrt{-5}$ 的复整数环内,整数 2、3、$-2+\sqrt{-5}$、$-2-\sqrt{-5}$、$1+\sqrt{-5}$、$1-\sqrt{-5}$ 都是不可约的.
14. 证明在形如 $a+b\sqrt{-5}$ 的复整数环内,主理想(2)等于 A^2, 这里 A 是理想$(2,1+\sqrt{-5})$.

割圆方程问题

15. 求 18 阶的循环群的 6 阶循环子群的陪集.
16. 用高斯的方法解割圆方程:$x^6+x^5+x^4+x^3+x^2+x^1+1=0$.
17. 在高斯解方程 $x^{19}-1=0$ 的例子中,证明 α_1,α_2 和 α_4 是三次方程 $x^3+x^2-6x-7=0$ 的根(提示:用这个方程的系数和根的对称函数之间的关系).
18. 在高斯解方程 $x^{19}-1=0$ 的例子中,证明 β_1,β_8 和 β_7 是三次方程 $x^3-\alpha_1x^2+(\alpha_2+\alpha_4)x-2-\alpha_2=0$ 的根,这里的 α 和 β 是原文中的 α 和 β.
19. 在高斯解方程 $x^{19}-1=0$ 的例子中,证明 r 和 r^{18} 都是 $x^2-\beta_1x+1=0$ 的根.这里的 r 和 β 是原文中的 r 和 β.

群论和数域的问题

注:其中的某些问题需要群论和伽罗瓦理论.

20. 两个等价的二次型有相同的判别式.
21. 用卡尔达诺的公式计算方程 $x^3+6x=6$ 在有理数域上的群 G,证明这个群有一个正规子群 H,并且 H 和 H 在 G 中的指数都是素的.
22. 写出 5 个 8 阶的群.
23. 通过找 3 次的模 5 不可约同余类来创造一个阶为 5^3 的域.
24. 比较韦伯的群的定义和现在的标准定义,证明它们是等价的.
25. 比较韦伯的域的定义和现在的标准定义.韦伯的某些公设是不是可以由其他的公设证明?

符号代数学的问题

26. 证明哈密顿关于数偶(α,β)的运算法则反映了复数 $\alpha+\beta\mathrm{i}$ 相似的运算法则.
27. 试构造一个形如 $\alpha+\beta\mathrm{i}+\gamma\mathrm{j}$ 的三元数的乘法,它满足哈密顿的关于有理数乘法的法则.即,这个乘法必需满足结合律、交换律、对加法是可分配的,并有惟一除法和模的乘法则.
28. 假设 $\alpha=3+4\mathrm{i}+7\mathrm{j}+\mathrm{k},\beta=2-3\mathrm{i}+\mathrm{j}-\mathrm{k}$ 是四元数,计算 $\alpha\beta$ 和 α/β.
29. 四元数 $\alpha=a+b\mathrm{i}+c\mathrm{j}+d\mathrm{k}$ 的模定义为 $|\alpha|=a^2+b^2+c^2+d^2$.证明 $|\alpha\beta|=|\alpha||\beta|$.
30. 把含有 3 个逻辑变量的一般形式的函数 $f(x,y,z)$ 展成形为 $x'y'z'$ 的多项式,比如,用 x' 代表 x 或者 $\bar{x}$.用此展开 $V=x-yz$.
31. 解释文中剩下的三个布尔等式 $x\bar{y}z=0,x\overline{yz}=0,\bar{x}yz=0$,其中 x 表示干净的野兽,y 代表分蹄的,z 代表反刍的.

矩阵代数问题

32. 证明如果变换 $x=\alpha x'+\beta y',y=\gamma x'+\delta y'$,其中 $\alpha\delta-\beta\gamma=1$,把二次型 $F=ax^2+2bxy+cy^2$ 转变成 $F'=a'x'^2+2b'x'y'+c'y'^2$,那么存在一个相同型的逆变换把后者变为前者.
33. 证明如果两个矩阵的乘积是 0 矩阵,那么至少有一个矩阵的行列式为 0.

34.具体证明凯莱-哈密顿定理:矩阵 A 满足特征方程 $\det(A-\lambda I)=0$,这里 A 是2×2矩阵.

35.证明矩阵

$$L=\begin{pmatrix}\frac{a+Y}{X} & \frac{b}{X}\\ \frac{c}{X} & \frac{d+Y}{X}\end{pmatrix}$$

是矩阵 $M=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ 的平方根.其中 $X=\sqrt{a+d+2\sqrt{ad-bc}}$,$Y=\sqrt{ad-bc}$.

36.上题的 2×2 矩阵 M 在什么条件下存在平方根,存在多少?

37.具体找出任意 3×3 矩阵 M 的平方根 L,利用 M 满足它的特征方程,以及 $M=L^2$ 的已知条件.

38.具体确定与 2×2 的矩阵 $M=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ 可交换的矩阵,并证明它们依赖于 M 的特征多项式的根的个数.

39.利用达朗贝尔方法,确定微分方程组

$$\frac{d^2y_i}{dt^2}+\sum_{k=1}^{3}a_{ik}y_k=0\quad(i=1,2,3)$$

的解.并证明为什么要使方程的解有物理意义必需矩阵(a_{ik})的三个特征值为相异的正实数.

40.用柯西的方法,找一个正交变换把二次型 $2x^2+6xy+5y^2$ 转换成平方的和或者差.

线性方程组的解的问题

41.解线性方程组:

$$\begin{aligned}2u+\ v+2x+\ y+\ 3x&=0\\ 5u+3v-4x+3y-\ 6z&=0\\ u+\ v-8x+\ y-12z&=0.\end{aligned}$$

首先确定系数矩阵的最大的非零行列式的阶.

42.确定上题中系数矩阵的秩,并找其解集的一个基.

43.利用上题的结论找一个与练习 41 的方程组相伴随的线性组,并找一组基,证明这个方程组与另外一个方程组相伴随,这个组与 41 的组同解.

讨论题

44.找一份韦伯 1895 年的《代数读本》(Lehrbuch der Algebra),并把它与现在的课本相比较.

45.设计一节课,通过:(a) 有限集上的置换概念;(b) 二次型的复合;(c) 模素数 p 的剩余类,引入群的概念.

46.克罗内克通过构造把类引入了代数数域,戴德金通过复数的子域把类引入代数数域,试比较这两种方法的优点和缺点.

47.比较德摩根关于代数的法则和韦伯关于域的公理.

48.设计一节课来解释负数,试通过皮科克的等价型和哈密顿的正数偶两种方法,那种方法较好,为什么?

49.中学生是如何"理解" 负数的?他们是否理解负数倍的负数就是正数?这种理解有必要吗?

50.设计一节课用哈密顿的有序数偶方法解释复数.

51.根据凯莱在 1858 年处理矩阵的方法,设计一节课,讲出矩阵运算的基本步骤.

文献和注解

代数史方面的总参考文献包括第 14 章中提到的三篇,即,由 Navý、Wussing 和 van der Waerden 写的书.其他有用的著作包括:Michael J. Crowe 的 *A History of Vector Analysis*: *The Evolution of the Idea of a Vectorial System*(Notre Dame:

University of Notre Dame Press, 1967) 和 Harold Edwards 的另两本书: *Fermat's Last Theorem. A Genetic Introduction to Number Theory*(New York: Springer-Verlag, 1977) 和 Galois Theory(New York: Springer-Verlag, 1984). 关于群论历史出色的综述是 Israel Kleiner 的"The Evolution of Group Theory: A Brief Survey", *Mathematics Magazine* 59(1986), 198 - 215. 还有几篇关于英国代数方面出色的论文,其中包括 Joan Richards 的"The Art and Science of British Algebra: A Study in the Perception of Mathematical Truth", *Historia Mathematica* 7(1980), 343 - 365; Helena Pycior, "George Peacock and the British Origins of Symbolical Algebra," *Historia Mathematica* 8(1981), 23 - 45, 以及 Ernest Nagel 的"'Impossible Numbers': A Chapter in the History of Modern Logic," *Historia of Ideas* 3(1935), 429 - 474. 上述文章中包含了比本文中更加详细的关于英国代数发展的内容. 书中关于矩阵论的历史来自 Thomas Hawkins 的三篇文章: "Cauchy and the Spectral Theory of Matrices," *Historia Mathematics* 2(1975), 1 - 29; "Another Look at Cayley and the Theory of Matrices," *Archives Internationales d'Histoire des Sciences* 26(1977), 82 - 112; 以及"Weierstrass and the Theory of Matrices," *Archive for History of Exact Sciences* 17(1977), 119 - 163. 阅读这些文章会得到矩阵论和相关数学领域发展的详细图像. 还有一篇更简短更基础的关于矩阵论发展的文章是 R.W.Feldmann 的"History of Elementary Matrix Theory," *Mathematics Teacher* 55(1962), 482 - 484, 589 - 590, 657 - 659, 以及 56(1963), 37 - 38, 101 - 102, 163 - 164.

1. Ernest Kummer, "De numeris complexis, qui radicibus unitatis et numeris integris realibus constant," *Journal de mathématiques pures et appliquées* 12(1847), 185 - 212, 重印于 Kummer's *Collected Papers*(New York: Springer-Verlag, 1975), vol. 1, 165 - 192, p. 182.
2. Michael J. Crowe, *Vector Analysis*, p. 29.
3. Carl Friedrich Gauss, *Disquistiones Arithmeticae*, Arthur A. Clarke 译(New York: Springer-Verlag, 1986), p. 1. 值得费心细读全书.
4. 同上, p. 35.
5. 一本高斯的新近的传记, 包括了他的著作目录和二次要文献概述; 这是 W.K.Bühler, *Gauss: A Biographical Study*(New York: Springer-Verlag, 1981).
6. 从高斯的 *Untersuchungen über höhere arithmetic*(New York: Chelsea, 1965), p. 540. 翻译而来. 此书由 *H. Maser* 对各种高斯关于"高等算术"的论文的德语翻译组成,最早在 1889 年出版.
7. Quoted in Edwards, *Fermant's Last Theorem*, p. 61. 两篇关于苏菲 · 杰曼的新近文章是 J.H.Sampson, "Sophie Germain and the Theory of Numbers." *Archive for History of Exact Sciences* 41(1991), 157 - 161 和 Amy Dahan Dalmédico, "Sophie Germain," *Scientific American*(December 1991), 117 - 122.
8. 关于 1847 年 5 月 1 日巴黎学术会议更详细的故事叙述可见于 Edwards, *Fermat's Last Theorem*, pp. 76 - 80.
9. 对除法和库默关于费马大定理更详细的讨论也可见 H.M.Edwards, "The Genesis of Ideal theory," *Archive for History of Exact Sciences* 23(1980), 321 - 378.
10. Richard Dedkind, "Sur la Théorie des Nombres entiers algébriques," *Bulletin des Sciences mathématiques et astronomiques* 11(1877)1 - 121, 它的部分内容重印于 *Mathematische Werke*, vol. 3, 262 - 296, p. 268.
11. 同上, p. 280.
12. 同上, p. 268. 关于理想概念的创造的更多内容见 H.M.Edwards, "Dedeking's Invention of Ideals," in Esther Philips, ed., *Studies in the History of Mathematics*(Washington: MAA, 1987), 8 - 20.
13. Gauss, *Disquisitiones*, p. 445.
14. 同上, p. 459.
15. 对阿贝尔的生平见 Oystein Ore, *Niels Henrik Abel: Mathematician Extraordinary*(New York: Chelsea, 1974).
16. 见 Van der Waerden, *History of Algebra*, pp. 85 - 88 中证明的细节.
17. 引自 Wussing, *Abstract Group Concept*, p. 98.
18. 同上, p. 100.
19. Galois, "Memoir on the Conditions for Solvability of Equations by Radicals," Harold Edwards 译, 收录于 Edwards, *Galois Theory*, p. 102.

20. 关于伽罗瓦命题和它们证明的详细讨论见 Edwards, *Galois Theory*. 在 H.M.Edwads, "A Note on Galois Theory," *Archive for History of Exact Sciences* 41(1991),163 - 169,中有进一步的澄清.

21. 一本现代最好的关于伽罗瓦生平的文章是 T.Rothman 的"Genius and Biographers:the Fictionalization of Evariste Galois," *American Mathematical Monthly* 89(1982),84 - 106.小说体的传记是 Leopold Infeld, *Whom the Gods Love: the Story of Evariste Galois*(New York:Whittlesey House,1948).一本关于决斗的新故事的新书是 Laura Toti Rigatolli, *Evariste Galois*,1811—1832,trans.by John Denton (Basel:Birkhäuser,1996).

22. 引自 Rothman,"Fictionalization of Evariste Galois,"p.97.

23. R.Bourgne and J.P.Azra.eds., *Ecrits et Mémoires Mathématiques d'Evariste Galois*(Paris:Gauthier-Villars,1962),p.185.

24. 同上,p.175.

25. Camille Jordan, *Traité des substitutions et des équations algébriques*(Paries:Gauthier-Villars,1870),sec.54.

26. 同上,sec.402.对伽罗瓦理论历史的更多的信息是 B.Melvin Kiernan, "The Development of Galois Theory from Lagrange to Artin ," *Archive for History of Exact Sciences* 8(1971),40 - 154.

27. Gauss, *Disquisitiones*, pp.264 - 265.

28. 同上,p.366.

29. 引自 Wussing, *Abstract Group Concept*, p.64.

30. Arthur Cayley,"On the Theory of Groups, as depending on the Symbolic Equation $\theta^n = 1$," *Philosophical Magazine*(4)7, 40 - 47,p.41. 此文也可见于 Cayley, *The Collected Mathematical Papers*(Cambridge:Cambridge University Press, 1889—1897),vol.2,123 - 130.

31. Arthur Cayley,"On the Theory of Groups," *Proceedings of the London Mathematical Society* 9(1878),126 - 133, p.127,此文重印于 Cayley, *Collected Mathematical Papers*, vol.10,324 - 330.

32. Arthur Cayley,"On the Theory of Groups," *American Journal of Mathematics* 1(1878),50 - 52,p.52.此文重印于 Cayley, *Collected Mathematical Papers*, vol.10,401 - 403.

33. 同上.

34. 引自 Wussing, *Abstract Group Concept*, p.235.

35. Walter Dyck,"Gruppentheoretische Studien," *Mathematische Annalen* 20(1882),1 - 44,p.1.对此文的进一步讨论见 Wussing, *Abstract Group Concept*, p.240 和 van der Waerden, *History of Algebra*, p.152.

36. 同上,p.12.

37. Heinrich Weber,"Beweis des Satzes,dass jede eigentlich primitive quadratische Form unendich viele Primzahlen fähig ist," *Mathematische Annalen* 20(1882),301 - 329.p.302,韦伯的工作也在 Wussing 和 van der Waerden 的书中有讨论.

38. Heinrich Weber,"Die allgenmeinen Grundlagen des Galios'schen Gleichungstheorie," *Mathematische Annalen* 43(1893), 521 - 549,p.524.

39. 引自 Kurt-R.Biermann. "Kronecker," *Dictionary of Scientific Biography*(New York:Scribners,1970—1980),vol.7, 505 - 509.

40. Richard Dedekind, Supplement XI to Drichlet, *Vorlesungen über Zahletheorie*(Braunschweig:Vieweg und Sohn,1893)p. 452.索引用的第四版,但在第二和第三版中也有这些资料.

41. William Rouse Ball, *The Origin and History of the Mathematical Tripos*(Cambridge:Cambridge University Press,1880),p.195

42. George Peacock, *A Treatise on Algebra, reprint edition*(New York:Scripta Mathematica,1940),vol.2,p.59.

43. 引自 Pycior,"George Peacock,"p.35.

44. Peacock, *Treatise on Algebra*, p.453.

45. 同上,p.449.

46. 引自 Abraham Arcavi 和 Maxim Bruckheimer,"The didactical De Morgan:a selection of Augustus de Morgan's thoughts on teaching and learning mathematics," *For the Learning of Mathematics* 9(1989),34 - 39.此文由德摩根的讲解他致力于更好的数学教学的著作中的许多摘录构成的.

47. Augustus de Morgan, *Trigonometry and Double Algebra*(London:Taylor, 1849)pp.92 - 93. 在 Nagel 的"Impossible

Numbers" 中被引用,pp.185 – 186.

48. Augustus de Morgan, "On the Foundation of Algebra, No. II," *Transactions of the Cambridge Philosophical Society* 7(1839—1842),287 – 300,p.287.

49. De Morgan, "On the Foundations of Algebra, No. II," pp.288 – 289.

50. Willam Rowan Hamilton, "Theory of Conjugate Functions or Algebraic Couples; with a Preliminary and Elementary Essay on Algebra as the Science of Pure Time," *Transactions of the Royal Irish Academy* 17(1837),293 – 422,重印于 Hamilton, *Mathematical Papers*(Cambridge: Cambridge University Press,1967)vol.3,3 – 96,p.4. 哈密顿工作的更多信息见 Jerold Mathrews, "Willam Rowan Hamilton's Paper of 1837 on the Arithmetization of Analysis," *Archieve for History of Exact Sciences* 19(1978),177 – 200 和 Thomas Hankins, "Algebra of Pure Time; William Rowan Hamilton and the Foundations of Algebra ," in P.J. Mackamer and R.G. Turnbull, eds., *Motion and Time, Space and Matter: Interrelations in the History and Philosophy of Science*(Columbus: Ohio State University Press, 1976),327 – 359.

51. Thomas L. Hankins, *Sir William Rowan Hamilton*(Baltimore: Johns Hopkins University Press,1980),p.343. 此传记是对哈密顿不仅在数学上而且在物理上工作的很好的研究.

52. Hamilton, *Mathematical Papers*, vol.3, p.10.

53. 同上,p.83.

54. 同上,p.96.

55. 引自 Crowe, *Vector Analysis*, p.27.

56. 同上,p.32.

57. Clerk Maxwell, *Treatise on Electricity and Magnetism*(London: Oxford University Press,1873),pp.9 – 10.

58. George Boole, *An Investigation of the Laws of Thought*(New York: Dover,1958),p.1. 关于布尔的更多信息是 D. MacHale, *George Boole: His Life and Work*(Dublin: Boole Press,1985).

59. 同上,p.27.

60. 同上,p.49.

61. 引自 Hawkins, "Another Look at Cayley," p.86.

62. James Joseph Sylvester, "On a New Class of Theorems," *Philosophical Magazine* (3)37 (1850),363 – 370, 重印于 *Collected Mathematical Papers*(Cambridge: Cambridge University Press,1904—1912)vol.1,145 – 151,p.150.

63. Arthur Cayley, "A Memoir on the Theory of Matrices," *Philosophical Transactions of the Royal Society of London* 148(1858), reprinted in Cayley, *Collected Mathematical Papers*, vol.2,475 – 496,p.476.

64. 同上,p.481.

65. 同上,p.483.

66. 同上.

67. 摘自 Charles L. Dodgson, *An Elementary Treatise on Determinants with their application to Simultaneous Linear Equations and Algebraical Geometry*(London: Macmillan,1867).p.50.

68. Georg Frobenius, "Über homogene totale Differentialgleichungen," Journal für die reine und angewandte Mathematik 86(1879),1 – 19,p.1. This paper is reprinted in Georg Frobenius, *Gesammelte Abhandlungen*(Berlin: Springer-Verlag, 1968), vol.1,435 – 453.

19世纪代数发展概览

1765—1822	鲍罗·鲁菲尼(Paolo Ruffini)	5次多项式方程
1776—1831	苏菲·杰曼(Sophie Germain)	数论
1777—1855	卡尔·弗里德里希·高斯(Carl Friedrich Gauss)	数论和割圆多项式
1789—1857	奥古斯丁·路易斯·柯西(Augustin-Louis Cauchy)	置换、行列式、特征值
1791—1858	乔治·皮科克(George Peacock)	符号代数
1802—1829	尼尔斯亨里克·阿贝尔(Niels Henrik Abel)	5次多项式方程
1805—1859	狄里克雷(Peter Lejeune-Dirichlet)	数论
1805—1865	威廉·罗万·哈密顿(William Rowan Hamilton)	复数、四元数
1806—1871	奥古斯塔斯·德摩根(Augustus De Morgan)	符号代数
1809—1882	约瑟夫·刘维尔(Joseph Liouville)	数论
1810—1893	恩斯特·库默尔(Ernst Kummer)	因式分解、因子
1811—1832	埃瓦里斯特·伽罗瓦(Evariste Galois)	方程的群、有限域
1814—1897	詹姆斯·约瑟夫·西尔维斯特(James Joseph Sylvester)	矩阵
1814—1848	皮埃尔·万泽尔(Pierre Wantzel)	作图问题
1815—1864	乔治·布尔(George Boole)	逻辑
1821—1895	亚瑟·凯莱(Arthur Cayley)	抽象群、矩阵
1823—1852	费尔迪南·哥特候德·爱森斯坦(Ferdinand Gotthold Eisenstein)	二次型、矩阵
1823—1891	奥波德·克罗内克(Leopold Kronecker)	抽象群、域
1826—1883	亨利·J.S.史密斯(Henry J.S.Smith)	线性方程组的解
1831—1879	詹姆斯·克拉克·麦克斯韦(James Clerk Maxwell)	四元数、向量
1831—1901	彼特·古日里·泰特(Peter Guthrie Tait)	四元数
1831—1916	理查德·戴德金(Richard Dedekind)	理想、域
1832—1898	查理斯·道吉森(Charles Dodgson)	线性方程组的解
1838—1922	卡米尔·约当(Camille Jordan)	群、线性变换
1839—1903	约西阿·魏拉尔德·吉布斯(Josiah Willard Gibbs)	向量
1842—1913	海因里希·韦伯(Heinrich Weber)	群、域
1849—1917	乔治·弗罗贝尼马斯(Georg Frobenius)	矩阵理论、方程组
1850—1925	奥列维·亥维塞(Oliver Heaviside)	向量
1852—1939	费尔迪南德·林德曼(Ferdinand Lindemann)	π的超越性
1856—1934	瓦尔特·迪克(Walter Dyck)	群论

第 16 章　19 世纪的分析

> 在前面提到过的柯西的著作《综合工科学校的分析教程》(Cours d'analyse de l'École Polytechnique) 中人们发现了下述定理:"如果级数 $u_0 + u_1 + u_2 + u_3 + \cdots$ 的不同项都是同一个单变量 x 的函数,并且在使该级数收敛的 x 的一个给定值的邻域中这些项是此变量的连续函数,则此级数的和 s 在这个给定值的邻域中也是 x 的一个连续函数."但是,据我看来,此定理确实有不成立的情形.例如级数 $\sin x - \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x - \cdots$ 对于 x 的值 $(2m+1)\pi$ 是不连续的,这里的 m 是整数.众所周知,有许多具有类似性质的级数.
>
> (N.H. 阿贝尔,摘自"关于级数 $1 + \frac{m}{1}x + \frac{m}{1}\frac{m-1}{2}x^2 + \frac{m}{1}\frac{m-1}{2}\cdot\frac{m-2}{3}x^3 + \cdots$ 的研究",1826)[1]

在 1858 年秋,苏黎世理工专科学校的教授理查德·戴德金被安排去作关于微分运算原理的讲座,这是以前还没有人讲过的内容.尽管在初等教程中所采用的对一些基本概念的传统几何方法具有教学上的价值,在准备这些讲座时,他认定微积分中那些与函数极限有关的部分仍然没有"科学的"基础.因此他决定集中精力去对实数概念的算术定义建立一个基础.1858 年 11 月 24 日,戴德金达到了他的目标,之后很快便将他的结果告诉了一个朋友和他的一些最好的学生.但是因为他觉得他的表述还不是完全自如的,故而在 1872 年之前他并没有发表他的"戴德金分割"思想.

在 18 世纪末,随着法国大革命重建整个欧州大陆数学教育的浪潮,并随着数学家不断增长的对教学的而非研究的需要,便产生了对应该如何把数学思想讲述给学生的不断增加的关切,随之而来的是对"严格性"的不断增长的关切.回想一下拉格朗日曾试图把所有微积分建立在幂级数概念的基础之上.虽然像其他一些人一样,拉克鲁瓦(Lacroix)运用拉格朗日的方法写了一本微积分的教程,但不久就发现并非所有的函数都可以用这种级数来表示.

奥古斯汀·路易斯·柯西,这位 19 世纪最多产的数学家,首先把微积分建立在极限概念的基础

上,这极类似于今天的情形.虽说在很久以前,甚至于牛顿就已经讨论过极限的概念,柯西却是第一个把趋向于一个特定值的函数这一有点模糊的概念转换成了算术语句的人,人们用它证明了极限的存在.柯西用他的极限概念来定义连续性(指的是现代意义上的连续性)和序列的收敛性,这里所说的序列既是数字的也是函数的.柯西的收敛概念发表于1821年,但在实质上早在1817年和1782年就已分别由捷克数学家伯恩哈德·波尔查诺和葡萄牙数学家库尼亚(José Anastácio da Cunha)阐述过.可惜后两人的著作出现在欧洲的偏远角落,没有引起法国和德国数学中心的重视,甚至连读都没有读到过.因此现今概念和发展源于柯西的工作.

柯西的一个重要结果是说,一个连续函数的无穷级数,如果它的和存在,则此和函数是连续的;但这个结论是错误的.反例早在1826年就发现了,它是现在被称为傅里叶级数的与正弦和余弦函数有关的级数.虽然在18世纪中叶丹尼尔·伯努利曾简要地考虑过这些级数,但第一次详尽地研究却是约瑟夫·傅里叶,它出现在19世纪初期他的关于热传导的著作中.傅里叶的工作激发了P.L.狄利克雷去更详细地研究函数的概念,也激发了B·黎曼去发展出现今称作黎曼积分的概念.

柯西和波尔查诺工作中的一些没有解决的问题以及由柯西的错误定理中产生的不连续点的研究使得许多数学家在世纪后半叶去考虑实数系统的结构.特别地,戴德金和康托尔各自发展出一套由有理数出发去构造实数的方法,从而开始了对无限集合的详细研究.

在他的微积分教程中,柯西把积分定义为一种和的极限而不像18世纪流行的那样,把它定义为反导数.1840年代他把这个积分概念推广到复数区域上,发展出了许多今天在复变函数初级教程中出现的重要概念,其中包括了留数的思想.这些思想在后来数十年中由黎曼进一步发展和推广了.

因为在复区域上的积分可以想像成在实二维平面上的积分,柯西也就能给出今天称作为格林定理的定理,它把沿一条闭曲线上的积分与此曲线所包围的区域上的二重积分联系起来了.把区域上的一个积分与在此区域边界上的积分联系起来的类似定理是由米哈依尔·奥斯特罗格拉茨基和威廉·汤姆逊发现的.这些定理现在被称作散度定理和斯托克斯定理,它们很快便被应用到物理学中如电磁学这样的领域.

本章所涉及的分析的最后一个方面是概率和统计在19世纪所取得的进展.我们要考虑勒让德和高斯在最小二乘法中的进展以及对正态误差曲线的高斯推导.在对拉普拉斯的对过去概率论思想的综合工作作简短考察之后,我们还要考虑凯泰勒特(Quetelet)的工作以及统计学的英国学派在发展这一学科某些基本工具方面的进展.

16.1 分析的严谨性

1799年拉克鲁瓦取代了拉格朗日在综合工科学校的位置.在此两年前,他的三卷本著作《论微分与积分》(Traité du calcul différentiel et du calcul intégral)的第一卷出版了.拉克鲁瓦把这本著作写成为自牛顿和莱布尼茨时代以来所发展出的微积分方法的综述.因此,拉克鲁瓦表述的不仅仅是拉格朗日的观点,即函数$f(x)$的导数是f的泰勒级数一阶项的系数,而且还按照达朗贝尔的方式将dy/dx定义为一种极限,也按欧拉的方式看作是无穷小量的比率.拉克鲁瓦为此书的广泛包容性感到自豪,并且希望这门学科的真正的形而上学会在各种不同方法的共同之处被发现.

然而,为了他在巴黎的教学,拉克鲁瓦把这三卷教程写成了一个缩减的一卷本,书名是《微积分导论》(Traité élémentaire du calcul différentiel et du calcul intégral),这是一本持续流行的教科书,这一点可由1802年到1881年间出现了九版这个事实得到印证.在这本著作中,拉克鲁瓦一开始就决

定把微分运算建立在一类极限的概念上,这类极限是在他确定微商的极限过程中定义出来的.因此,他指出,如果 $u = ax^2, u_1 = a(x+h)^2$,则 $2ax$ 是“比值 $(u_1 - u)/h$ 的**极限**,或者说是当量 h 变小时这个比值均衡地趋向的那个值,而且它可按我们选取的程度来靠近这个值.”[2] 在计算了许多其它值的极限后,拉克鲁瓦解释说,事实上,“微分运算就是求函数和它所依赖的变量变化时产生的增量间比值的极限.”[3] 于是拉克鲁瓦,欧拉和拉格朗日,他们都没有试图在书的开始就把微分运算启发为切线的斜率.微分运算是“分析”的一部分,并不需要几何的动机或图解.切线只不过是这种运算的一个应用,而且在书的第 7 节就有关于它的讨论,这一节的标题是“关于微分运算在曲线论中的应用”.

尽管拉克鲁瓦在《导论》中决定用极限来作为讲述微分运算的开端,但他却很快地转到建立函数的泰勒级数上去了.像拉格朗日一样,他相信所有函数,除了可能在孤立点外,都可以表示为级数.于是他运用这种泰勒级数表示去建立了各种超越函数的微分公式,发展了确定极大值和极小值的方法,甚至对多变量函数也这样做.在后一种情形中,他纠正了欧拉在确定双变元函数取极值的条件上的错误.事实上,他证明了如函数 $u(x,y)$ 在一点的两个一阶偏导数均为零时,函数在这点具有极值的一个充分条件是,在这点成立



图 16.1 法国邮票中的柯西.

$$\frac{\partial^2 u}{\partial x^2} \cdot \frac{\partial^2 u}{\partial y^2} > \left[\frac{\partial^2 u}{\partial x \partial y}\right]^2.$$

人物小传

奥古斯汀·路易斯·柯西(1789—1857)(Augustin-Louis Cauchy)

虽然柯西(图 16.1)是 19 世纪最多产的数学家,但他绝不是一个容易相处的人.正像阿贝尔在 1826 年访问巴黎时写给他一个朋友的信中所说的:“尽管他是目前最了解数学应该是怎样的数学家,但没法与他好好相处 … 柯西是狂热的天主教徒而且对教义非常偏执,对一个数学家而言,这是极其奇怪的.就其他方面来说,他是目前(在巴黎)惟一一个在纯数学领域中工作的人.”[4] 他于法国大革命的年代诞生在首都,接受了优良的古典式教育,而后从 1805 年到 1807 年在综合工科学校学习工程.当他从 1810 年到 1813 年作为工程师参与拿破仑政府的许多军事工程的工作时,他表现出对纯数学的强烈兴趣,这使得拉普拉斯和拉格朗日鼓励他离开工程工作,在他们的帮助下在综合工科学校得到了一个数学职位,多年之后也在法兰西学院工作.由于他在分析方面的教科书的出现,他成了法国数学界中最受尊敬的成员之一.他写了那么多的数学文章,以致巴黎科学院的杂志强行限制了任何个人发表文章的数量.柯西创办了自己的杂志以摆脱这种限制.

当 1830 年的七月革命推翻了最后一个波旁皇帝时,作为一个狂热的保守份子柯西拒绝宣誓效忠于新国王,并自我流放到了意大利,而后又去了布拉格.1838 年他回到了巴黎但没有回到他的教学岗位,这种情况一直延续到 1848 的革命,这时取消了效忠宣誓的要求.或许正是由于他的政治观点的缘故,法国政府一直没有发行纪念他的邮票,直到他诞生二百周年时才第一次发行.

拉格朗日的幂级数方法对剑桥分析学会那些有改革意识的成员也颇具吸引力,他们之中包括了乔治·皮柯克,查理·巴贝奇和约翰·黑尔舍(1792—1871),他们把拉克鲁瓦的《导论》翻译成英文,作为在剑桥使用的一本分析教程.译者们对拉克鲁瓦“用达朗贝尔的极限方法取代了更为正确

和自然的拉格朗日方法"感到失望,以致他们在译本中加上了注解,要读者们运用拉格朗日的方法而不要用极限方法.

不管拉格朗日方法在英国多么具有吸引力,回到了法国的柯西却发现它缺乏"严谨".事实上对那些他相信是没有牢靠的基础的代数表达式,特别是无穷长的表达式的处理方法,柯西是不满意的.涉及到这些表达式的方程只对某些值是正确的,对这些值这个无穷级数是收敛的.特别,柯西发现函数 $f(x) = e^{-x^2} + e^{-(1/x^2)}$ 的泰勒级数并不收敛于这个函数.因此,由于从 1813 年起他正在综合工科学校教课,柯西开始了对微积分基础的彻底完全的重新思考.1821 年,在他的许多同事的催促下,他发表了他的《皇家综合工科学校分析教程》,在此书中他引进了微积分基础中的新方法.我们要研讨一下出现在这本书和随后于 1823 年发表的《在皇家综合工科学校课程中关于无穷小计算的概要》(Résumé des Leçons données a l'École Royale Polytechnique sur le Calcul Infinitésmal)的分析内容中的关于极限、连续性、收敛性、导数及积分的柯西的思想,因为正是这些在巴黎使用的教科书对于这个世纪的余下年代提供了微积分教程的标准模式.

补遗 16.1

什么是极限?

莱布尼茨(1684):如果任何一个连续变迁以一个极限为终结,那么就能够形成一种普遍的推理,它也能适用于最终的极限.

牛顿(1687):逐渐变小的量之间的最终比值…(是)极限,即数量比值无限减小却总是收敛于它;它们比任何事先给定的差值更接近地趋向于它,但永远不超过也不达到它,直到这些量减到无穷小.

麦克劳林(1742):$2x+o$ 与 a 的比率当 o 减小时连续地减小,并当 o 是任意实增量时总是大于 $2x$ 与 a 的比,而这显示了它连续趋向于 $2x$ 与 a 的比,并以其为极限.

达朗贝尔(1754):比值 $[a:2y+z]$ 总是小于 $a:2y$,但是 z 越小,这个比值就越大,并且由于人们可选取任意小的 z,比值 $a:2y+z$ 就可按我们希望的那样靠近比值 $a:2y$.因此 $a:2y$ 是比值 $a:2y+z$ 的极限.

拉克鲁瓦(1806):比值 $(\mu_1-\mu)/h$ 的极限…是这样一个值,当量 h 减小时这个比值按比例趋向于它并且可按我们作出的选择来趋近于它.

柯西(1821):如果赋予同一变量的连续不断的一系列数值使其无限地趋向于一个固定的值,使得最终它们与固定值的差按人们所希望的那样小,则后者称为所有其它数值的极限.

16.1.1 极限

柯西的极限定义出现在《分析教程》的开篇部分:"如果赋予同一变量的连续不断的一系列数值使其无限地趋向于一个固定的值,使得最终它们与固定值的差按人们所希望地那样小,则后者称为所有其它数值的**极限**".[6] 作为例子,柯西指出一个无理数是趋向于它的各种不同分数的极限(补遗 16.1).他还定义一个**无穷小量**是一个以零为极限的变量.请注意,柯西没有定义现代的概念 $\lim\limits_{x\to a} f(x) = b$,因为那种概念涉及到两个不同的变量.他似乎完全抑制了自变量的作用.另外,看起来似乎柯西的极限定义与达朗贝尔的没有多大差别.然而要看出柯西的定义文字上的意思究意是什么,并去发现它与他以前的人们给出的定义之间的差别,就有必要去考察一下他运用这个定义去

证明的有关极限的一些特殊结果.事实上,柯西不只是处理了非独立的和独立的变量,而且还用不等式的语言把他的陈述作了算术的翻译.作为一个例子,考虑柯西对下面定理的证明:

定理　如果对 x 的增加值,差 $f(x+1)-f(x)$ 收敛于某个极限 k,则分式 $f(x)/x$ 也同时收敛于同一极限.

柯西一开始便将定理的假设转换成了算术陈述:给定一个任意小的值 ε,人们可以找到一个数 h,使得当 $x \geqslant h$ 时有 $k-\varepsilon < f(x+1)-f(x) < k+\varepsilon$.现在他能够把这种转换了的陈述用到他的证明中去了.因为对 $i=1,2,\cdots,n$ 的每个差 $f(h+i)-f(h+i-1)$ 满足这个不等式,故它们的算术平均值

$$\frac{f(h+n)-f(h)}{n}$$

也满足这个不等式.从而推导出

$$\frac{f(h+n)-f(h)}{n}=k+\alpha,$$

其中 $-\varepsilon<\alpha<\varepsilon$,或者令 $x=h+n$,便有

$$\frac{f(x)-f(h)}{x-h}=k+\alpha.$$

但是由此得到 $f(x)=f(h)+(x-h)(k+\alpha)$,或者

$$\frac{f(x)}{x}=\frac{f(h)}{x}+\left(1-\frac{h}{x}\right)(k+\alpha).$$

因为 h 是固定的,柯西得出当 x 变大时 $f(x)/x$ 趋向于 $k+\alpha$,其中 $-\varepsilon<\alpha<\varepsilon$.因为 ε 是任意的,定理的结论成立.柯西还证明 $k=\pm\infty$ 时定理成立,从而利用这个结果可以得出以下结论:例如,当 x 变大时 $\log x/x$ 收敛于 0 而 $a^x/x(a>1)$ 的极限是 ∞.

16.1.2　连续性

给出极限定义之后,柯西现在便能定义至关重要的连续性概念了.回想一下,对欧拉而言,一个连续函数是一个用单一表达式表示的函数,而一个不连续函数则是在它的定义域的不同部分有不同表达式的函数.柯西意识到这种定义是自相矛盾的.例如,函数 $f(x)$ 在 x 为正值时等于 x,而在负值时等于 $-x$ 时,按欧拉的定义是不连续的.另一方面,人们可以用一个单一的解析表达式

$$f(x)=\frac{2}{\pi}\int_0^{\infty}\frac{x^2}{t^2+x^2}\mathrm{d}t$$

把它表示出来,故而 $f(x)$ 又是连续的.连续曲线的几何观点通常被理解为一条没有断裂的曲线,而柯西要寻找到一个分析的定义来对函数表达出这种观点.拉格朗日早先曾试图在特殊情况作出这样的定义,这种情形是数“在 0 点连续”并在此处取值为 0:“我们总能找到一个对应于小于任何一个给定数的纵坐标的横坐标 h;从而使所有小于 h 的值也对应于小于给定数的纵坐标.”[8]

补遗 16.2

连续性的定义

欧拉(1748):一条连续曲线是一条其性质可由 x 的一个单一函数来表达的曲线.如果一条曲线具有这种性质,即对它的不同部分……需要用 x 的不同函数来表达……那么我们就称这样的曲线为不连续.

波尔查诺(1817):对 x 的所有在某个界限以内或以外的所有值,函数 $f(x)$ 按连续性法则变化是指……如果(当)x 是某个这种值时,差 $f(x+\omega)-f(x)$ 在 ω 被取作任意小时可以小于任何预先给定的量.

柯西(1821):称函数 $f(x)$ 是在变量 x 的两个给定值之间的这个变量的连续函数是指对 x 在这些界限之间的每个值,差 $f(x+\alpha)-f(x)$ 的(绝对)值随着 α 无限地变小.

狄利克雷(1837):设想 a 和 b 为两个固定值,而作为一个变量的 x 可以逐次取 a 与 b 之间的所有值.现在,如果对每个 x 对应了单个的有限值 y,其方式是当 x 连续地穿过从 a 到 b 的区间时,$y=f(x)$ 也逐渐地变化,那么 y 就被称作 x 在此区间上的一个连续函数.

海涅(1872):称一个函数 $f(x)$ 在一个特定值 $x=X$ 处连续,是指如果对每个给定的量 ε,不管多小,存在一个正数 η_0.它具有如下性质:没有小于 η_0 的正数 η 使 $f(x\pm\eta)-f(x)$ 的绝对值超过 ε.如果 $f(x)$ 对 $x=a$ 和 $x=b$ 之间的每一个值 $x=X$,包括 $x=a$ 和 $x=b$,是连续的,则说 $f(x)$ 从 $x=a$ 到 $x=b$ 连续.

读了拉格朗日的著作后,柯西推广了拉格朗日的思想并给出了他自己的新定义:"称函数 $f(x)$ 是在变量 x 的两个给定值之间的这个变量的**连续函数**是指对 x 在这些界限之间的每个值,差 $f(x+\alpha)-f(x)$ 的(绝对)数值随 α 而无限地变小,(见补遗 16.2).换句话说,如果在这些值之间,变量的一个无限小增量总能产生出函数自身的一个无限小增量,则函数 $f(x)$ 对于在给定界限之间的 x 仍然保持着连续性."[9] 注意,柯西既提出了一个算术的定义也提出了一个用更为熟悉的无穷小量的语言叙述的定义.但是因为柯西已经以极限的语言定义了这些量,所以这两个定义的意思是一样的.柯西以证明例如 $\sin x$(在任意区间)是连续的来展示如何运用他的定义.因为

$$\sin(x+\alpha)-\sin x = 2\sin\frac{1}{2}\alpha\cdot\cos\left(x+\frac{1}{2}\alpha\right),$$

而右端清楚地显示随 α 无限地减小.

有意思的是一位捷克的数学家伯恩哈德·波尔查诺(1781—1848),他也熟悉拉格朗日的工作,并早在大约四年前已经给出了连续性的一个定义,而这个定义实际上与柯西的是一样的.作为他严格证明**介值定理**的计划的一部分,波尔查诺需要给出使定理成立的那类数的一个清晰定义.介值定理是说"如果未知量的两个值(代入一个连续函数 $f(x)$ 时)得到相反的符号,则在其间必定至少有方程($f(x)=0$)的一个实根存在."[10] 留意到其他一些人利用非数学的概念,如像时间,运动等,已给出了连续性的定义,波尔查诺也给出了他称之为的"正确定义":"对 x 的所有在某个界限以内或以外的所有值,函数 $f(x)$ 按连续性法则变化是指……如果(当)x 是某个这种值时,差 $f(x+\omega)-f(x)$ 在 ω 被取作任意小时可以小于任何预先给定的量."[11] 尽管这两个定义非常类似,但并没有令人信服的证据表明柯西在做出自己的定义时曾读到过波尔查诺的著作.这两个人都感兴趣于给出一个定义并由此使某个"显然的"结果能得到证明;他们想出了本质上相同的



图 16.2 捷克斯洛伐克邮票上的波尔查诺.

主意.

> 伯恩哈德 · 波尔查诺(1781—1848)(Bernhard Bolzano)
>
> 人物小传
>
> 波尔查诺(图 16.2)在他的家乡布拉格的大学学习过数学,哲学和物理.1805 年他被任命担任这里的宗教哲学方面的一个重要职位,这个职位是按奥地利皇帝的命令设立的,用以对抗紧随法国大革命后席卷整个欧洲的新的启蒙潮流.但是波尔查诺不是一个天主教复兴的特别同情者,而且在他的讲座中表达了他自己在宗教方面的开明观点.终于在 1819 年他被解除了这个职位并因异教徒的嫌疑被置于警察的监管之下,但是,与此同时,他的哲学训练把他吸引到关于分析的基础这种问题上,通过对直观的极限和连续观念给出新的定义和证明,他得以满意地解决这些问题.

当然,从现代的观点来看,这两个人都没有定义在一点的连续性.他们定义的都是在一个区间上的连续性.但是,人们似乎可以把每个定义都解读为在区间的每个点上定义了连续性.即给出一个特别的点 x 和一个 $\varepsilon > 0$,可以找到一个 $\delta > 0$,使得当 $\alpha < \delta$ 时有 $|f(x+\alpha) - f(x)| < \varepsilon$.虽然如此,但我们在下一节会看到,柯西并不完全清楚 δ 依赖于什么量.由于缺乏这种明晰性,导致他得出了一个不正确的结果.

16.1.3　收敛性

柯西的级数收敛概念也有先行者.但正是他的定义才伴随有明晰的判别准则来实际验证收敛性,这些准则直到今天还在使用.柯西的定义出现在他的《分析教程》的第 6 章:"设 $s_n = \mu_0 + \mu_1 + \mu_2 + \cdots + \mu_{n-1}$ 为(一个级数的)前 n 项的和,n 代表任意整数.如果对 n 的增加值,和数 s_n 无限趋向于一个极限 s,则称此级数是**收敛的**,上述的极限被称为此级数的**和**.相反地,如果当 n 无限增大时,和 s_n 不趋近于任何一个固定的极限,则它是**发散的**,且没有和."[12] 为了把定义解释清楚,柯西陈述了现在所谓的收敛性的"柯西判别准则".他认识到为了使级数收敛则每个单项 μ_n 必须减小至零.这个条件不是充分的.如果各种和 $\mu_n + \mu_{n+1}, \mu_n + \mu_{n+1} + \mu_{n+2}, \mu_n + \mu_{n+1} + \mu_{n+2} + \mu_{n+3}, \cdots$"从第一个起到人们希望取的任何数目为止,其绝对值总小于任何事先给定的界限"[13],此时才能保证级数的收敛性.柯西没有提供这一充分条件的证明,这是因为没有实数的某种算术定义便不可能证明它.但是他确实提供了例子.他证明了几何级数 $1 + x + x^2 + x^3 + \cdots$ 当 $|x| < 1$ 时收敛,因为和数 $x^n + x^{n+1}, x^n + x^{n+1} + x^{x+2}, \cdots$ 各自等于 $x^n \dfrac{1-x^2}{1-x}, x^n \dfrac{1-x^3}{1-x}, \cdots$,它们总介于 x^n 和 $\dfrac{x^n}{1-x}$ 之间,后面的两个值当 n 增加时都收敛于 0.

在他的柯西判别准则出现之前就已超过柯西的不仅有波尔查诺而且还有葡萄牙学者约瑟 · 阿纳斯塔西奥 · 达 · 库尼亚(1744—1787),这方面的内容包含在他的《数学原理》(Principios Mathematicos)之中;这是一本包容广泛的教科书,从算术的基本知识一直讲到变分计算,并且从 1782 年开始便逐章地出版.达 · 库尼亚版本的收敛性和柯西判别准则是这样叙述的:"数学家们说,一个**收敛**级数是那样的级数,它的每项都以相似的方法确定,即每项由前面的诸项决定,所依方式应能使此系列总能连贯进行,而最终人们可以忽略不计他们希望加到已经写下或表示出的那些和上去的任意多项之和,这不会引起明显的误差."[14] 达 · 库尼亚甚至像柯西一样运用这个准则去证明几何级数的收敛性.虽然达 · 库尼亚的著作在 1811 年被翻译成了法文,但可惜它显然没有引起什么注意也没有什么影响.

人物小传	约瑟·阿纳斯塔西奥·达·库尼亚(1744—1787)(José Anastácio Da Cunha)
	达·库尼亚在里斯本受教育并在1762年的法国和西班牙入侵葡萄牙时成了一名军官.在从事弹道学的研究时,他于1769年写了一本关于这门学科的专著,它分析了这方面的各式各样的指导手册.他的著作引起了波姆巴候爵的注意,而后者是约瑟一世皇帝的有权势的大臣.当时国王已能削弱宗教法庭的权力并把基督教士从权位上赶走.1773年波姆巴把达·库尼亚安排到重新组建的柯印布拉大学,出任一个几何方面的重要职位.但1777年国王逝世后,波姆巴失去了权力而他的被保护者,已获有自由思想者名声的库尼亚则被宗教法庭逮捕并控以异教罪.他于1781年得到赦免,并在里斯本的一个为穷苦儿童教育建立的学校中,作为一名数学教授度过了他的余生.

波尔查诺的著作同样发表在欧洲的偏远角落,也很少有即时的影响.但在波尔查诺的著作中证明了柯西判别准则蕴含了最小上界原理,这个原理最终被看成是实数系统的界定性质之一.波尔查诺的收敛性定义和他的柯西判别准则的陈述包含在下面定理之中:

定理　如果一系列的量 $F_1(x), F_2(x), \cdots, F_n(x), \cdots$(这里的每个 $F_i(x)$ 可以想成是代表了一个级数前 i 项之和),具有如下性质,即如果取 n 足够大,则序列的第 n 项 $F_n(x)$ 与每个后面的项 $F_{n+r}(x)$ 的差仍然小于任何一个给定的量;这里的后面项不管取得离前面的有多远都如此,那么总存在一个而且只有一个常量,使此序列的项趋向于它,而且只要此序列延续到足够远,它们便能按所期望的那样靠近它.[15]

波尔查诺对极限惟一性的证明是直截了当的,但他对于对每个 x 存在级数收敛的数 $X(x)$ 的证明却是错误的,因为像柯西一样,波尔查诺无法定义一个任意的实数 X.虽然如此,他确实指出了如何在任给精度 d 之内来确定这个 X.如果 n 取得充分大,使得对每个 r, $F_{n+r}(x)$ 与 $F_n(x)$ 之差小于 d,则 $F_n(x)$ 就是所要的 X 的近似值.

波尔查诺现在也可以证明实数的最小上界性质了.

定理　如果一个性质 M 不属于变量 x 的所有值但是属于小于某个数 μ 的所有值,则总存在一个量 U,它是所有那种数中最大的一个,而那些数中每一个都使得所有小于它的 x 具有性质 M.

波尔查诺对所有具有性质 M 的数的最小上界 U 的存在性证明涉及到构造一个级数,而对这个级数可以应用收敛判别准则.因为 M 不对每个 x 都成立,则必存在量 $V = \mu + D$,使得对小于 V 的每个 x, M 成立的这个断言为假.现在波尔查诺考虑对应于每个正整数 m 的量 $V_m = \mu + D/2^m$.如果对所有 m,断言 M 对所有小于 V_m 的所有 x 都成立是假的,那么 μ 本身必是所要的最小上界.另一方面,设若 M 对所有小于 $\mu + D/2^m$ 的 x 都成立,但对小于 $\mu + D/2^{m-1}$ 的 x 不全成立.这两个量的差为 $D/2^m$.下一步波尔查诺便对区间 $[\mu + D/2^m, \mu + D/2^{m-1}]$ 运用二分法技巧,确定出一个最小的整数 n 使得对所有小于 $\mu + D/2^m + D/2^{m+n}$ 的 x, M 成立而对所有小于 $\mu + D/2^m + D/2^{m+n-1}$ 的 x, M 不全成立.继续这种做法,波尔查诺便构造出一个序列 $\mu, \mu + D/2^m, \mu + D/2^m + D/2^{m+n}, \cdots$,它满足他的柯西判别准则,因此必定收敛于一个值 U,他很容易便能证明这个值满足定理中的条件.(波尔查诺的证明在1860年代被魏尔斯特拉斯作了小小的改动以用来证明现在所称的波尔查诺-魏尔斯特拉斯定理.该定理说,任给一个实数的有界无限集合 S,则存在一个实数 r 使得它的每个邻域中总存在 S 的其他点.)

由最小上界原理容易推出介值定理.设 $f(\alpha) < 0$,而 $f(\beta) > 0$.不失一般性,我们也可假设对所有 $x < \alpha$ 有 $f(x) < 0$,那么 $f(x) < 0$ 这个性质 M 并不对全部 x 成立,但是对所有小于某个数 $\mu =$

$\alpha + \omega$ 的所有 x 成立,其中 $\omega < \beta - \alpha$(因为假定 f 是连续的).由此存在数 U,它是使对所有 $x < U$,$f(x) < 0$ 成立的最大的数.证明 $f(U)$ 既不能是正的也不能是负的,这是直截了当的事,因此定理得证.

对这个同一结果的柯西的证明则没有使用柯西判别准则.作为替换,这个证明隐含地依赖于数系的另一个公理,即任一有界单调序列有极限,同时还明显地依赖于他早先曾证明过的一个结果,即如果 $f(x)$ 是连续的并有序列 $\{a_i\}$ 收敛于 a,则序列 $\{f(a_i)\}$ 收敛于 $f(a)$.柯西的证明步骤来源于拉格朗日和早先的数学家们在逼近多项式方程 $f(x) = 0$ 的解时所用的标准逼近过程.于是,如果 $f(\alpha) < 0$ 与 $f(\beta) > 0$,则令 $h = \beta - \alpha$,m 为一任意正整数.柯西考虑了 $f\left(\alpha + \frac{ih}{m}\right)$ 的符号,其中 $i = 1,2,\cdots$. 因为有一对邻接的数 α_1 和 β_1,满足 $\alpha < \alpha_1 < \beta_1 < \beta$ 及 $f(\alpha_1) < 0$,$f(\beta_1) > 0$,下一步柯西则把长度为 h/m 的区间$[\alpha_1,\beta_1]$分为长度为 h/m^2 的子区间并重复同一推断.一直下去,他便得到了一个递增序列 $\alpha,\alpha_1,\alpha_2,\cdots$ 和一个递降序列 $\beta,\beta_1,\beta_2,\cdots$,每个序列必收敛于同一极限 a.由此得出两上序列 $f(\alpha),f(\alpha_1),f(\alpha_2),\cdots$ 和 $f(\beta),f(\beta_1),f(\beta_2),\cdots$,它们都收敛于 $f(a)$.由于第一个序列的每个值都是负的而第二个序列的每个值都是正的,故必然有 $f(a) = 0$,从而定理得证.然而值得注意的是,这个证明出现在柯西的教科书的附录中,而书的正文中他只给了一个几何的推断,即看出曲线 $y = f(x)$ 和 $y = 0$ 必然互相穿越.人们只能推测说,柯西实际上只在课堂上讲了这个方法.

在给出了级数收敛性的柯西判别准则后,在他的教科书中柯西又发展出各种判别法,利用它们可以证明一些特别情形下的收敛性;书中是从关于正项级数的判别开始的,比如说是 $\mu_0 + \mu_1 + \mu_2 + \cdots$. 比较判别法是说,如果所给级数逐项以一个收敛级数的项为界,则所给级数本身便收敛;柯西使用了这个判别法而没有作任何特别的评说.他最常用的是与公比小于 1 的几何级数的比较,而后者的收敛性由柯西用柯西判别准则一开始就证明了.事实上,他利用比较判别法去证明了他的许多其他判别法的有效性.举例说,柯西证明了根式判别法,即如果数$\sqrt[n]{\mu_n}$ 的极限是个小于 1 的数 k,则此级数收敛.选取数 U 使$k < U < 1$,柯西注意到,当 n 充分大时有$\sqrt[n]{\mu_n} < U$或者写成$\mu_n < U^n$.然后用与收敛几何级数 $1 + U + U^2 + U^3 + \cdots$ 比较得出所给级数也收敛.类似地,如果这些根式的极限是个大于 1 的数,则所给级数发散.

柯西以类似的方法证明了比值判别法:如果当 n 增大时比值μ_{n+1}/μ_n 收敛于一个固定极限k,则当 $k < 1$ 时级数 μ_n 收敛,当 $k > 1$ 时它发散.

对于既有正项也有负项的级数,柯西以绝对收敛的思想来处理(尽管没有用这个术语),把根式判别和比值判别修改到适合这种情况,也证明了交错级数判别法,还指出了如何计算两个收敛级数的和与积.特别是他指出了如何找出一个幂级数的收敛区间.虽说关于级数的一些结果在先前已经知道,柯西却是第一个人把它们组织成一个协调一致的理论,这使得他和其他一些人能够把这套理论推广到复数和函数的级数上去.

在《分析教程》中有个有意思的结果,虽说如阿贝尔在 1826 年注意到的那样,这是个错误的结果;它是说:

定理 6.1.1 当级数$\left\{\sum_{n=0}^{\infty}\mu_n\right\}$的不同项都是同一变量 x 的函数,并对此变量在一个使级数收敛的特定值的邻域内这些函数是连续的,那么此级数的和 s 在此特定值的邻域中也是 x 的连续函数.

对此结果柯西的"证明" 十分简单.我们以柯西所用的语句来展示他的推断,然后把这些语句

转换成现代的符号.记 s_n 为些级数前 n 项的和,s 为整个级数的和,柯西以 r_n 表示余项 $s - s_n$.(这里 s,s_n 和 r_n 均是 x 的函数.)为证明 s 的连续性他需要证明 x 的一个无穷小增量导致 $s(x)$ 的一个无穷小增量,就是说,给定 $\varepsilon > 0$,$\exists\,\delta$ 使得

$$\forall a, |a| < \delta \Rightarrow |s(x+a) - s(x)| < \varepsilon. \tag{16.1}$$

虽然早先的某些证明中柯西实际计算过 δ 的适当值,而在这个情形中,他试图进行的只是使用任意无穷小量作一个推断.因此他写下的是,x 的一个无穷小增量 x 导致了 $s_n(x)$ 的一个无穷小增量,因为后者对所有 n 均为连续,或者表示为 $\exists\,\delta$ 使得

$$\forall a, |a| < \delta \Rightarrow |s_n(x+a) - s_n(x)| < \varepsilon. \tag{16.2}$$

下一步,因为级数对任一 x 都收敛,那么 r_n 对 n 足够大时其本身也为无穷小,并对 x 的无穷小增量也是这样的,或者表示 $\exists\,N$,使得

$$\forall n, n > N \Rightarrow |r_n(x)| < \varepsilon \quad \text{和} \quad |r_n(x+a)| < \varepsilon. \tag{16.3}$$

因为 s 的增量是 s_n 和 r_n 增量的和,柯西得出结论说,这个增量也是无穷小的,因此 s 本身连续.我们把它写成

$$\begin{aligned}|s(x+a) - s(x)| &= |s_n(x+a) + r_n(x+a) - s_n(x) - r_n(x)| \\ &\leqslant |s_n(x+a) - s_n(x)| + |r_n(x+a)| + |r_n(x)| \\ &\leqslant \varepsilon + \varepsilon + \varepsilon = 3\varepsilon.\end{aligned} \tag{16.4}$$

柯西的失察之处在于(16.2)中的 δ 依赖于 ε, x, n,而(16.3)中的 N 依赖于 ε, x 和 a.除非我们知道存在某个数 N,对于(16.3)中的所有满足 $|a| < \delta$ 的 a 均成立,否则我们便不能断言(16.4)为真(或说(16.1)为真).柯西的用无穷小增量的推断掩盖了所涉及到的各个量之间所需要的关系.要使证明有效必须用到一致收敛性的观念,即是说,人们需要另外的假定:数 N 的选取能够与 x 无关,至少也要在某个固定的区间中做到这点.我们将在 16.1.8 中回到这个问题上来.

16.1.4　导　数

在柯西的《分析教程》中提供了函数和级数的基本思想的一个处理办法.在他 1823 年的教科书《在皇家综合工科学校课程中关于无穷小计算的概要》中柯西用他关于极限的新思想研究了导数和积分,这是无穷小计算的两个基本概念.

在此书的开始用了以前书中的同样的连续性定义之后,柯西在第三课中进行了函数导数的定义,他把它称之为**导函数**(fonction deriveé),即为 $[f(x+i) - f(x)]/i$ 当 i 趋向极限 0 时的极限,只要这个极限是存在的.恰如他在连续性的定义中所做的那样,柯西定义了在一个区间上的导数概念,实际上就是函数在上面连续的那个区间.他注意到这个极限在 x 的每个值上有一个确定的值,从而是那个变量的一个新函数,他用拉格朗日记号 $f'(x)$ 来表示它.对定义自身而言,虽然可以像欧拉的著作中那样想成是无穷小差的商这种表达式,但他却更直接地取自于拉格朗日的书《解析函数》中,在那里拉格朗日把它作为 f 的幂级数展开部分证明了 $f(x+i) = f(x) + if'(x) + iV$,其中 V 是某个随 i 趋向零的函数.柯西便有可能把这个关于导数的定理转换成了导数的定义.然后他计算了许多初等函数的导数.例如,若 $f(x) = \sin x$,则定义中的商可简约为

$$\frac{\sin\,(1/2)^i}{(1/2)i}\cos(x + (1/2)i),$$

可以看出它的极限 $f'(x)$ 是 $\cos x$.

自然,柯西的导数计算没有什么新东西,柯西能证明的关于导数方面的定理也没有任何特别的

新意.拉格朗日从他自己的导数定义推导出了同样的结果.但是因为拉格朗日的导数定义建立在一个错误的假定之上,即任何函数都可以展开成幂级数,那么柯西的工作的意义就在于他明确地应用了导数的现代定义,通过他的极限定义转换成不等式的语言来证明定理.从后面要用到的观点看,这些结果中最重要的一个在第七课中:

定理　如果函数 $f(x)$ 在 $x = x_0$ 与 $x = X$ 之间连续,并且如果让 A 和 B 分别是导数 $f'(x)$ 在此区间中的最小和最大值,则有限差的比

$$\frac{f(X) - f(x_0)}{X - x_0}$$

必在 A 与 B 之间.[17]

柯西对此定理的证明首次使用了当今学生们如此熟悉的 δ 和 ε.柯西由选取 $\varepsilon > 0$ 开始,而后选取 δ 使得对所有 $|i| < \delta$ 的 i 和对区间 $[x_0, X]$ 中的 x 的任一值,不等式

$$f'(x) - \varepsilon < \frac{f(x+i) - f(x)}{i} < f'(x) + \varepsilon$$

成立.从柯西的作为一个极限的导数定义推导出这样的值是存在的.但是注意,柯西所用的事实是隐含在他对一个区间上的导数定义中而不是在一个点上的定义中,就是说,对给出的 ε,用一个 δ 对区间中每个 x 都起作用.不管怎样,下一步柯西在 x_0 与 $x_n = X$ 之间置入了 $n - 1$ 个新的值 $x_1 < x_2 < \cdots < x_{n-1}$,满足对每个 i 有 $x_i - x_{i-1} < \delta$,并把上述不等式应用到由相邻两个值定出的子区间上.这就得出对 $i = 1, 2, \cdots, n$ 有

$$A - \varepsilon < \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} < B + \varepsilon.$$

然后柯西利用一个代数结果得出结论说,分子的和除以分母的和也必定满足同一不等式,即

$$A - \varepsilon < \frac{f(X) - f(x_0)}{X - x_0} < B + \varepsilon.$$

由于这结果对每个 ε 都成立,从而得出了定理的结论.

作为此定理的直接推论,柯西推导出了导数的中值定理.设 $f'(x)$ 在给定的区间上连续,这个假定自然保证了它具有最小值 A 和最大值 B 的条件成立.柯西用函数的中值定理推导出 $f'(x)$ 取 A 与 B 之间的每个值,其中 x 在 x_0 与 X 之间.特别地,$f'(x)$ 可取定理中的那个值.因此存在一个值 θ,它介于 0 和 1 之间,使得

$$\frac{f(X) - f(x_0)}{X - x_0} = f'(x_0 + \theta(X - x_0)).$$

运用这个中值定理柯西便证明了:在一个区间上具有正导数的函数为递增的,具有负导数的函数是递降的,而且有零导数的函数是常数.

16.1.5　积　分

尽管用了他的极限新定义,柯西的导数处理还是紧密地与欧拉和拉格朗日著作中的处理相关联.另一方面,柯西对积分的处理却完全开辟了新的天地.回想在 18 世纪,积分被简单地定义为微分的逆.拉克鲁瓦甚至写道:“积分运算就是微分运算的逆,它的目标就是从微分系数出发提升到它们被导出来的函数.”[18] 虽然莱布尼茨制定出他的记号以提示人们积分是无穷小面积的无限和,但是在使用无限性中的固有问题使 18 世纪的数学家相信采用不定积分或者反导数的观念作为积分理论的基本概念是对的.自然人们还是认识到,要算出面积不能只用反导数还需用各种逼近技巧.

但正是柯西第一个把这些技巧当作基本的东西并依此构造出一个定积分的理论.

或许有许多条理由来说明柯西为什么不得不以和的极限而不是利用反导数来定义积分.首先,有许多种情况清楚表明,一条曲线下方的面积甚至当反导数在区间端点算不出来时也是有意义的,这种情形特别发生在某种逐段连续的函数,它已在傅里叶关于三角函数级数的书中被揭示出来.第二个理由是由将在16.3.2节中讨论的柯西的著作中充分阐述的,即是为了发展一套复函数的积分理论.最后,柯西或许在组织他在综合工科学校的讲稿材料时已经意识到,不能保证对每个函数都存在反导数.但对他选择用和来定义积分的理由柯西自己的解释是:它行得通.如他在1823年的一篇文章中所说,"对于我以这种样子来构想一个定积分(当作在积分符号下的微分表达式的无穷小值的和)似乎应该是按偏爱而采用的 … 因为它平等地适用于所有情形,甚至对那些一般说来不能够从$\int$符号下面的函数过渡到原函数的情形也可以."[19] 更进一步,他指出"当我们采用这种方式来构思定积分时,我们很容易证明,每当变量的两个界限为有限而且$\int$符号下的函数本身在包含于这两个值间的整个区间上保持有限和连续时,这样的积分便有一个惟一的和有限的值."[20]

在他的《概要》的第二部分中柯西阐述了以和给出的积分的严谨定义.大概是从欧拉和拉克鲁瓦关于定积分的近似计算的工作中取得了他的定义.但与其以逼近一块面积的方式来考虑这个方法还不如说大体从直观上了解了它的存在.柯西把逼近做成了一个定义.于是,设$f(x)$在$[x_0,X]$上连续,他取了x_0与$x_n = X$之间的$n-1$个新的中间值$x_1 < x_2 < \cdots < x_{n-1}$,并形成了和

$$S = (x_1 - x_0)f(x_0) + (x_2 - x_1)f(x_1) + \cdots + (X - x_{n-1})f(x_{n-1}).$$

柯西留意到,S依赖于n同时也依赖于所选取的特定值x_i.但是他写道:"重要的是要观察到,如果区间$[x_{i+1} - x_i]$变得非常小而数n非常大时,分割的方法对S的值只有察觉不出来的影响."[21]

为证明这个结果,柯西注意到如果以对原来的每个子区间进行细分的方式选取一个新的细分,那么相应的和S'可以被重写为下面的形式

$$S' = (x_1 - x_0)f(x_0 + \theta_0(x_1 - x_0)) + (x_2 - x_1)f(x_1 + \theta_1(x_2 - x_1)) + \cdots$$
$$+ (X - x_{n-1})f(x_{n-1} + \theta_{n-1}(X - x_{n-1})),$$

其中每个θ_i介于0与1之间.由连续性的定义,这个表达式可以重写成

$$\begin{aligned} S' &= (x_1 - x_0)[f(x_0) + \varepsilon_0] + (x_2 - x_1)[f(x_1) + \varepsilon_1] + \cdots + \\ &\quad (X - x_{n-1})[f(x_{n-1}) + \varepsilon_{n-1}] \\ &= S + (x_1 - x_0)\varepsilon_1 + (x_2 - x_1)\varepsilon_2 + \cdots + (X - x_{n-1})\varepsilon_{n-1} \\ &= S + (X - x_0)\varepsilon', \end{aligned}$$

其中ε'是介于所有ε_i中的最小值与最大值之间的值.然后柯西论证说,如果细分足够的小,则这些ε_i和ε'也将非常靠近零,那么加细的分割并不会对和数有可观的改变.给出两个充分细小的子分割时,人们可以取第三个分割来当作它们中每一个的子分割.对第三个的和值则可任意地靠近对原来两个分割中的每一个的和值.由此推出"如果我们以增加它们的数目来让(细分长度的)数值无限制地减小,最终S的值则明显地变为常数,或者换句话说,它将以达到某个极限而终结,而此极限惟一地依赖于函数$f(x)$的形式以及给予变量x的端值x_0和X.这个极限便是我们称为的定积分$\left(\text{写作}\int_{x_0}^{X} f(x)\,\mathrm{d}x\right)$."[22] 这个定义用了柯西收敛判别准则的一种推广形式,它推广到了并非以自然数为指标的序列.

以这种用和的极限定义的积分,柯西不难证明积分的中值定理,即

$$\int_{x_0}^{X} f(x)\mathrm{d}x = (X - x_0)f(x_0 + \theta(X - x_0)),$$

其中 $0 \leqslant \theta \leqslant 1$,也不难证明区间上的积分可加性定理.然后他轻而易举地证明了:

微积分基本定理　如果 $f(x)$ 在 $[x_0, X]$ 中连续,如果 $x \in [x_0, X]$,且

$$F(x) = \int_{x_0}^{x} f(x)\mathrm{d}x$$

则 $F'(x) = f(x)$.

为证明此定理柯西运用了中值定理和可加性得到

$$F(x + \alpha) - F(x) = M\int_{x}^{x+\alpha} f(x)\mathrm{d}x = \alpha f(x + \theta\alpha).$$

如果用 α 除以两端并取极限,便可由 $f(x)$ 的连续性得到结论.可以认为基本定理的这种版本是第一个符合现代严谨性标准的,因为在其中 $F(x)$ 第一次通过对定积分的存在性证明被清晰地定义了.

虽然对 $f(x)$ 的定积分存在性的柯西的原来的假定条件之一是 $f(x)$ 在积分区间上连续,但柯西也认识到甚至在某种程度上放宽这个假定条件他的定义仍有意义.因此他指出,如果在给定区间上 $f(x)$ 有有限多个不连续点,则积分仍然可以定义,这只要将区间在不连续点处剖开为子区间并以进一步的极限论证来定义积分就行了.类似地,如果 $f(x)$ 在 $(a, b]$ 上连续,柯西作的定义是

$$\int_{a}^{b} f(x)\mathrm{d}x = \lim_{\varepsilon \to 0}\int_{a+\varepsilon}^{b} f(x)\mathrm{d}x.$$

对函数在无穷区间上的积分柯西也作了类似的定义.

在他对满足 $\partial A/\partial y = \partial B/\partial x$ 的微分式 $A\mathrm{d}x + B\mathrm{d}y$ 的积分新方法的陈述中,柯西使用了他的积分定义的一个不同的推广.柯西指出所要的积分 $f(x, y)$ 可以在平面中由一个固定点 (x_0, y_0) 出发取定积分

$$f(x, y) = \int_{x_0}^{x} A(x, y)\mathrm{d}x + \int_{y_0}^{y} B(x_0, y)\mathrm{d}y$$

来定义.因为这个定义意味着 $\partial f/\partial x = A(x, y)$,并且

$$\begin{aligned}\frac{\partial f}{\partial y} &= \int_{x_0}^{x} \frac{\partial A}{\partial y}(x, y)\mathrm{d}x + B(x_0, y)\\ &= \int_{x_0}^{x} \frac{\partial B}{\partial x}(x, y)\mathrm{d}x + B(x_0, y)\\ &= B(x, y) - B(x_0, y) + B(x_0, y) = B(x, y),\end{aligned}$$

这个函数确实给出了问题的解答.

《分析教程》和《概要》是柯西在综合工科学校第一年课程的基础教材.在第二年里他对微分方程论作了详细的介绍.这个课程中大量涉及的是在 18 世纪已发展起来的解决这类问题的标准技巧,当然包括了上面所提到的定理.但是以他对分析严谨性新的关切,柯西需要确定一些条件,在这些条件下人们可以证明满足事先所给初始条件的 $y' = f(x, y)$ 解的存在性.在他证明中用到的逼近技巧是由给定的初始点 (x_0, y_0) 开始构造小的直线线段来逼近所希望的曲线,这种方法本质上是在 18 世纪已经用过的.但正是柯西,他能够运用柯西判别准则去证明,如果 $f(x, y)$ 和 $\partial f(x, y)/\partial y$ 都在包含 (x_0, y_0) 的一个平面区域内为有限、连续和有界,则这个逼近法便生成多边形,它们收敛于微分方程在 (x_0, y_0) 的包含于原区域中某邻域内的解曲线.

有一个与柯西处理微分方程有关的怪事.柯西从来没有发表过第二学年课程的报告,只是在最近,这些课程的前 13 讲的清样稿才为人所知.不清楚为什么这些讲稿在这里就停止了,但有证据表明柯西受到了学校校长的责备.他被告之,因为综合工科学校基本上是个工科学校,他应该用课堂时间来教授微分方程的应用而不是处理严谨性问题.柯西被迫遵照执行并宣布,他将不再给出完全严谨的证明了.因此他显然觉得他不能发表这种教材的讲稿,因为它们没有反映他自己的应如何处理这门学科的观念.

16.1.6 傅里叶级数和函数概念

对柯西的关于收敛性错误结果的阿贝尔的反例是一个傅里叶级数,即一个三角函数级数,它是欧拉和丹尼尔·伯努利在 18 世纪中叶曾争论过的那一类级数.约瑟夫·傅里叶(1768—1830) 于 19 世纪初在考虑热扩散问题时对它进行过详细研究.他先把它的著作在 1807 年提交给了法国科学院,而后经重写和扩充,成为他的《热的分析理论》(Théorie analytique de la chaleur),发表于 1822 年.傅里叶开始考虑的是一种温度分布的特殊情形,温度分布为 $v(t,x,y)$,t 为时间,分布在一个矩形薄片上,其正 x 方向无限长,y 方向的宽度 2,边 $x=0$ 维持着常温 1,而边 $y=\pm 1$ 保持温度 0.在对热流作了某些假定后,傅里叶能够证明 v 满足偏微分方程

$$\frac{\partial v}{\partial t}=\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}.$$

而后傅里叶在某种条件下解出了这个方程,这个条件是薄片的温度已达到了均衡,即 $\partial v/\partial t=0$.假定 $v=\varphi(x)\psi(y)$(分离变量法),他对每个变量微分两次得到 $\varphi''(x)\psi(y)+\varphi(x)\psi''(y)=0$ 或者

$$\frac{\varphi(x)}{\varphi''(x)}=-\frac{\psi(y)}{\psi''(y)}=A,$$

A 为某个常数.这些方程的显然的解是 $\varphi(x)=\alpha e^{mx}$,$\psi(x)=\beta\cos ny$,其中 $m^2=n^2=1/A$.物理的推断强令 m 为负值(否则对大的 x 值温度将趋向无穷大),故而傅里叶得出结论说原偏微分方程的一般解是形如 $v=ae^{-nx}\cos ny$ 这样一些函数的和.运用当 $y=\pm 1$ 时 $v=0$ 的边界条件傅里叶证明 n 必是 $\pi/2$ 的奇数倍,因此这个一般解由下面的无穷级数给出:

$$v=a_1e^{-\left(\frac{\pi x}{2}\right)}\cos\left(\frac{\pi y}{2}\right)+a_2e^{-(3\pi x/2)}\cos\left(\frac{3\pi y}{2}\right)+a_3e^{(-5\pi y/2)}\cos\left(\frac{5\pi y}{2}\right)+\cdots.$$

人物小传

让·巴布蒂斯塔·约瑟夫·傅里叶(1768—1830)(Jean Baptiste Joseph Fourier)

在九岁成为孤儿后,傅里叶便被置于他家乡奥舍尔的主教监护之下,此地距巴黎西南 90 英里.在当地一所军事学校中,他很快显示出了数学的天才.虽然他希望成为一位军队工程师,但因为他的出身并不高贵,他不能获得这样的职业.因而他取得了一个教书职位.随着大革命的爆发,傅里叶成了地方事务的显要人物.但是由于他对 1794 年恐怖事件受害者的保护,导致了他的被捕.幸运的是,在罗伯斯庇尔死后他被释放并在 1795 年被任命为综合工科学校的拉格朗日和蒙日的助手.三年之后,在拿破仑的埃及战役期间,他成了埃及学院的秘书,在这个位置上他能够对埃及的古迹进行广泛的研究.回到法国后,他在法国东南部的伊泽尔省当了 12 年的省长,并成功地完成了许多公众事业的改进.幸运的是,即便在拿破仑倒台之后,他仍然能够被选到重建的科学院,并在 1822 年成为它的终身秘书,用现代的词来说,即它的行政院长,生前他一直在此职位上.

为定出系数 a_i，傅里叶使用了另外的边界条件，即当 $x = 0$ 时 $v = 1$. 令 $\mu = \pi y/2$，这表明这些 a_i 满足方程

$$1 = a_1 \cos \mu + a_2 \cos 3\mu + a_3 \cos 5\mu + \cdots,$$

这是有无穷多个未知量的单个方程式，但傅里叶常以无限次微分它并每次都令 $\mu = 0$ 的方法把它变成无穷多个方程式. 注意到所解出的这些方程前面若干个已定出的形式，傅里叶便能确定 $a_1 = 4/\pi, a_2 = -4/3\pi, a_3 = 4/5\pi, \cdots$，因而解出了这个偏微分方程. 但以数学家们通常的性格，一旦原来的问题解决了，他便会开始考虑他的新型解的衍生物. 首先，他注意到这些系数的值蕴含了

$$\cos \mu - \frac{1}{3}\cos 3\mu + \frac{1}{5}\cos 5\mu - \cdots = \frac{\pi}{4},$$

其中 $\mu \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. 但这个级数显然对 $\mu = \frac{\pi}{2}$ 表示了 0 而对 $\mu \in \left(\frac{\pi}{2}, \frac{3}{2}\pi\right)$ 表示了 $-\frac{\pi}{4}$. 傅里叶意识到这个结果不会很快被读者们相信，但是，"由于这些结果看起来远离了微积分的通常结果，有必要对它们进行仔细地考虑并解释出它们的真实意义. 把方程式

$$y = \cos \mu - \frac{1}{3}\cos 3\mu + \frac{1}{5}\cos 5\mu - \frac{1}{7}\cos 7\mu + \cdots$$

考虑成一条曲线，其中 μ 为横坐标而 y 为纵坐标. 可以看见 … 这条曲线必定由分隔开的部分 aa, bb, cc, dd, … 组成，它们中每一条都平行于坐标轴且长度等于[π]. 这些平行线段被交错地置于轴的上方和下方，距离为[$\pi/4$]，并且被垂线 ab, cb, cd, ed, … 相连接，它们自身也是曲线的部分."[23] 换句话说，傅里叶把这个无穷的余弦级数考虑成图 16.3 中表示的"方形波". 对现代的读者来说，不清楚为什么傅里叶会画上垂直于横轴的线段，因为级数在 $\mu = k\pi/2$ 当 k 为奇数时总是 0. 然而，傅里叶在意识到级数的部分和总表示了一条没有断裂的曲线时，便以为无穷和应当由这样一条曲线来表示. 此曲线是否表示了现代意义上的"函数"或是否此级数代表了柯西意义下的连续函数这一类问题对于傅里叶的工作来说都不合适. 他感兴趣的是一个物理问题，或许以几何方式构想答案，在这些地方他可以画一条"连续"曲线而并不担心它是否代表了一个"函数".

下一步，傅里叶考察了由三角函数级数给出的各种函数表示，最一般的是形如 $c_0 + c_1 \cos x + c_2 \cos 2x + \cdots + d_1 \sin x + d_2 \sin 2x + \cdots$ 的级数. 但是通常傅里叶只限于或是正弦的或是余弦的级数. 另外，由于他要说服他的读者相信他的方法的有效性，他用了新的方法来确定原有的和其他的级数的系数. 例如，他指出，如果把函数 $(1/2)\pi f(x)$ 写成正弦的级数

$$\frac{1}{2}\pi f(x) = a_1 \sin x + a_2 \sin 2x + a_3 \sin 3x + \cdots,$$

人们可以轮流对每个 n 在等式两边乘以 $\sin nx\,dx$，并在区间[$0, \pi$]上积分. 因为

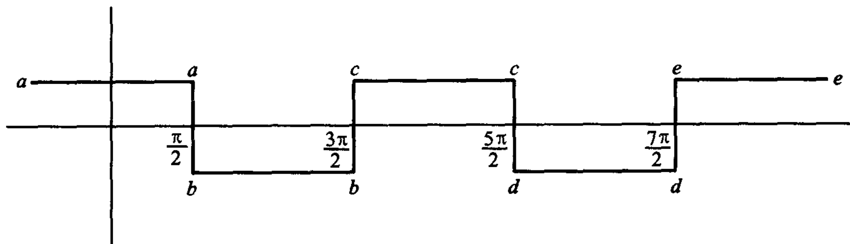


图 16.3 傅里叶的方形波：$y = \cos \mu - \frac{1}{3}\cos 3\mu + \frac{1}{5}\cos 5\mu - \frac{1}{7}\cos 7\mu + \cdots$.

$$\int_0^\pi \sin mx \sin nx\mathrm{d}x = \begin{cases}0, & \text{如果 } m \neq n;\\ \pi/2, & \text{如果 } m = n,\end{cases}$$

则得到

$$a_k = \int_0^\pi f(x)\sin kx\mathrm{d}x,$$

只要这个表示系数的积分有意义就行，这就是说，只要在 $f(x)\sin kx$ 下方的面积确有定义.

那么，什么类型的函数能够被这些三角函数级数表示呢？在开始回答这个问题时，傅里叶定义了它所用的词"函数"的意思："一般地，函数 $f(x)$ 代表了一系列的值或纵坐标，它们的每一个都是任意的. 在给出横坐标 x 的无穷多个值的同时，存在同等数量的纵坐标 $f(x)$. 所有这些都有实际的数值，或者为正，为负或为零. 我们并不假定这些纵坐标遵从于一个公共的法则；它们以任何一种随意的方式相互连接，给出它们中的每一个就好像它是单个的量一样."[24] 且不管这个听起来颇现代的定义，傅里叶从来就没有考虑过现今被称作的"任意函数". 他的例子表明他只想考虑逐段连续函数. 当然，傅里叶只断言级数表示了在一个特定的有限区间如像 $[0,\pi]$ 的内部给出的任意函数. 级数在端点的值可以很容易地单独计算，而正弦函数的周期性使人们可以把原来的函数几何地拓展到整条实直线上(见补遗 16.3).

补遗 16.3　函数是什么？

约翰·伯努利(1718)：我称一个由以任何方式从变量和从常数中构成的量为这个变量的函数.

欧拉(1748)：一个变量的函数是一个分析表达式，它按任何方式由变量，数和常量构成.

欧拉(1755)：当一些量按这样一种方式依赖于其他量，即当(后者)变化时(前者)本身也经历了变化，则称(前者)为(后者的)函数；这是一个内涵非常广泛的思想，它本身包括了所有由其他量来决定一个量的方式.

拉克鲁瓦(1810)：如果一个量的值依赖于一个或多个量，则称它为后者的函数，不管人们是否知道要使用什么运算来从后者达到前者.

傅里叶(1822)：一般地，函数 $f(x)$ 代表了一系列的值或纵坐标，它们的每一个都是任意的. 在给出横坐标 x 的无穷多个值的同时，存在同等数量的纵坐标 $f(x)$. 所有这些都有实际的数值，或者为正，为负或为零. 我们并不假定这些纵坐标遵从于一个公共的法则；它们以任何一种随意的方式相互连接，给出它们中每一个就好像它是单个的量一样.

海涅(1872)：一个变量 x 的单值函数是一个表达式，它由 x 的每个单个的有理或无理数值惟一确定.

戴德金(1888)：集合 S 上的一个函数 φ 是一个法则，按此法则，对 S 中每个确定的元素 s 有所属于的一个确定东西，称它为 s 的变换，并以 $\varphi(s)$ 表示.

在某些情形，傅里叶试图用三角恒等式把前 n 项的部分和重写为简明的形式，然后考虑当 n 增大时的极限来证明他的展开式实际代表了这个函数. 但是，总体上他相信他的以表示实面积的积分来明晰计算他所建议的一个任意函数展开式的系数是一个足以令人信服的论证，证明了这个展开式是有效的. 例如，他计算了后来由阿贝尔给出的

$$\frac{1}{2}x = \sin x - \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x - \frac{1}{4}\sin 4x + \cdots.$$

这个级数在 $[0,\pi/2)$ 上代表 $(1/2)x$，而在整条实直线上是图 16.4 的函数. 阿贝尔认识到这个函数不

仅违反了关于连续函数和的柯西结果,同时也说明傅里叶在证明傅里叶级数收敛于原来函数的尝试是不充分的.

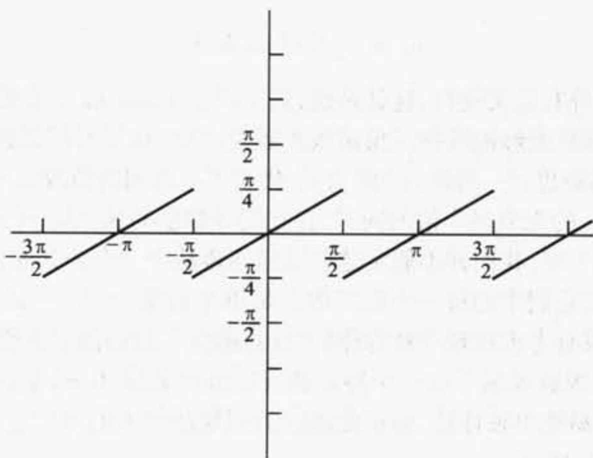


图 16.4　$y = \sin x - \frac{1}{2}\sin 2x + \frac{1}{3}\sin 3x - \frac{1}{4}\sin 4x + \cdots$ 的傅里叶的图形.

人物小传	古斯塔夫 · 彼得 · 勒热纳 · 狄利克雷(1805—1859)(Gustav Peter Lejeune-Dirichlet)
	由于在他青年时期,他家乡德国的数学教育水平低下,狄利克雷于 1822 年到了巴黎的巴黎学院学习.他在一位著名的法国将军家当了家庭教师,因而能够遇见许多最著名的法国知识界人物,其中包括约瑟夫 · 傅里叶,后来证明傅里叶对他的数学有极强的影响.狄利克雷于 1825 年返回德国,三年后被任命为柏林大学的教员,在此职位上他一直干了 27 年.他在柏林有非常重的教学负担,这部分由于他同时在军事科学院和大学中教课.因此,在 1855 年高斯逝世后,他愿意接受邀请到哥廷根,在这里他有了许多的研究时间.不幸的是,这在他 1859 年逝世前只持续了三年半.

柯西自己试图作出傅里叶在 1826 年的断言的新证明,但在那个证明中他假定了如果 $(u_k - v_k) \to 0, k \to \infty$ 和 $\sum u_k$ 收敛,则对于 $\sum v_k$ 也同样收敛.狄利克雷在 1829 年的一篇文章中注意到这个假设是错误的,它包含有反例

$$u_k = \frac{(-1)^k}{\sqrt{k}},\ v_k = \frac{(-1)^k}{\sqrt{k}} + \frac{1}{k},$$

显然其中第一个级数收敛而第二个则发散.然后,狄利克雷成功地攻克了这个傅里叶级数的收敛性问题,其证明是柯西的分析风格.狄利克雷不是试着去证明对"任意"函数的傅里叶级数收敛而是退而去寻求能保证收敛的那些函数的充分条件.特别地,他证明,如果函数 $f(x)$ 除去可能在有限个有限不连续点外在此区间连续并有界,而且在此区间中只有有限个转向值,则此傅里叶级数在 $(-\pi,\pi)$ 的每个 x 上收敛于 $\lim_{\varepsilon\to 0}\frac{1}{2}[f(x-\varepsilon)+f(x+\varepsilon)]$(在 f 连续时等于 $f(x)$).如果像傅里叶那样,把 $f(x)$ 在给定区间外解释为几何式的周期扩张,则此结果对端点也成立.狄利克雷之所以选

取这些条件是为了能够对所给函数与三角函数的积在某个区间上进行积分.但柯西对定积分的新规定只能保证对具有限个不连续点的函数的积分存在性.狄利克雷认识到把这个结果推广到在给定区间上具无限个不连续点函数的困难所在.事实上,他提供了一个函数作为例子,这个函数不满足他原来的条件,是一个处处不连续的函数:"$f(x)$ 在变量 x 取有理数时等于一个确定的常数 c,而当此变量为无理数时等于另一个常数 d.这个被定义的函数对每个 x 值具有有限而确定的值,但同时人们不能把它替换成级数,这只要看到那些在级数中的不同积分在此时都失去了任何意义就明白了."[25]

16.1.7 黎曼积分

在 1853 年,乔治·伯恩哈德·黎曼(1826—1866)曾尝试按柯西对积分 $\int_a^b f(x)\mathrm{d}x$ 的定义首先来准确决定哪些函数是可积的,用以推广狄利克雷的结果.事实上,他开始时便对此定义作了某种程度的改变.像柯西一样,他把区间 $[a,b]$ 分成 n 个子区间 $[x_{i-1},x_i]$, $i=1,2,\cdots,n$.令 $\delta_i=x_i-x_{i-1}$,他现在考虑了和

$$S=\sum_{i=1}^{n}\delta_i f(x_{i-1}+\epsilon_i\delta_i),$$

其中每个 ϵ_i 介于 0 和 1 之间.因为黎曼允许函数 f 的变量在适当的子区间中取任何值,所以这个和较之柯西的要一般得多.然后它定义积分为 S 所趋向的极限,只要这个极限存在而不管这些 δ_i 和 ϵ_i 是怎样取的.现在黎曼问了一个柯西不曾问过的问题:在什么情形下一个函数可积而又在什么情形下不可积?柯西自己只指出过某一类函数是可积的,但没有试图找出所有这类函数.另一方面,黎曼列出了一个有限函数 $f(x)$ 为可积的充分必要条件:"如果随着所有的 δ 量的无限减小,那些使函数 $f(x)$ 在上面的变差大于一个给定的 σ 量的区间的总长度 s 最终总变为无穷小,那么当所有 δ 量变为无穷小时和 S 收敛,"[26] 反之亦然.(一个函数在一个区间上的**变差**是这个函数在那区间上最大值与最小值的差.)作为不满足可积性柯西准则却是黎曼可积的例子是一个定义在 $[0,1]$ 上的函数,它是由黎曼给出的:

$$f(x)=\sum_{n=1}^{\infty}\frac{\varphi(nx)}{n^2},$$

其中 $\varphi(x)$ 等于 x 减去最靠近它的整数,要是存在两个同等靠近的整数则定义为 0.最后证明除了在无穷多个点 $x=p/2n$ 外 f 是连续的,这里的 p 与 n 互素.但是因为靠近这样一个点的 f 的变差等于 $\pi^2/8n^2$,那么使靠近它们时变差大于任意 $\sigma>0$ 的点的个数为有限,从而此函数满足黎曼可积性准则.

用现在所证明可积的一类新函数,黎曼便能把狄利克雷的结果推广到傅里叶级数的收敛性方面去.但是与其决定一个函数能保证使它的级数收敛的充分条件,黎曼认为还不如去攻克相反的问题,即从一个可由三角函数表示的函数开始,尝试着决定这个表示对于这个函数的行为会有什么结果.依此方式黎曼能找出许多种函数,它们都能由三角函数级数表示,但这从来就没有把整个问题回答到使他满意的程度.大概这就是黎曼从没有发表过这方面材料的手稿的原因.

16.1.8 一致收敛性

狄利克雷和黎曼的工作十分清楚地表明了一个傅里叶级数可以表示一个不连续的函数,从而对连续函数级数和的柯西定理必须作出修订.这由许多位数学家所完成,但正是卡尔·魏尔斯特拉

斯认识到该如何像在柯西原来的结论那样保证和函数在整个区间上连续. 1850 年代开始的在柏林大学的讲课过程中,魏尔斯特拉斯在数项级数的收敛性与函数项级数的收敛性之间作了仔细的区分,而柯西对此区分作了不恰当的解释. 然后他有了可能识别出函数收敛性的一个至关重要的性质,即在一个区间上的一致收敛性. 一个无穷级数 $\sum u_n(x)$ 在 $[a,b]$ 中**一致收敛**是指当给出任意的 $\epsilon > 0$, 存在一个 N(依赖于 ϵ) 使得对每个 $n > N$ 和对区间 $[a,b]$ 中每个 x 都满足 $|r_n(x)| < \epsilon$.

人物小传

卡尔 · 魏尔斯特拉斯(1815—1897)(Karl Weierstrass)

魏尔斯特拉斯生于德国的威斯特法伦,在 1834 年进入波恩大学,在其父亲的催促下学习公共财经和行政,其目的是在普鲁士行政部门谋得一个职位. 他对数学的天生的爱好再加上对于波恩小酒馆的伙伴情谊的兴趣妨碍他在他所拟定的领域中表现出色. 1838 年他没有取得学位便离开了波恩. 为了维持生活,他不得不为取得一纸教师证书再去学习;从 1841 年开始的 14 年间他在各种高级中学中教授像数学、物理、德语、植物学、地理、历史、体育和书法这些课程. 在一系列光辉的数学论文发表在克雷尔的《杂志》上后,他于 1854 年被授予了哥尼斯堡大学的荣誉博士学位并最后在 1856 年接受了柏林大学的副教授的任命,同时又取得了柏林工学院的教授头衔. 虽然健康问题使他只能坐着讲课并由一个高级学生替他写黑板,但他清晰的授课立刻使他赢得了在欧洲范围内的名声,每年都有数百的学生来听他的课. 1861 年他与库默尔一起把数学讨论班引进到柏林大学,这是使这个大学成为在 19 世纪后期的全世界在纯数学方面的首要大学的另一个因素.

在给出了魏尔斯特拉斯的定义后,那么纠正柯西对要求级数的一致收敛性的那种情形的错误证明便是十分简单的事了. 但是这个定义还有一个更深层的影响. 魏尔斯特拉斯不只是在他的定义中把某些量如何依赖于其他量弄得完全清楚,而且他也完成了从使用像"无穷小"这类词句中摆脱出来的转变. 自此以后, 所有涉及这种思想的定义都完全被算术地给出. 例如, 爱德华 · 海涅(1821—1881),一位哈雷的教授,他用了许多时间在柏林与魏尔斯特拉斯讨论数学;他不但在 1872 年的一篇论文中给出了在一个点上的连续定义,而且还把柯西的在一个区间上的连续定义重新改写为下面的样子:"一个函数被称作 …… 从 $x = a$ 到 $x = b$ 为**一致连续**的,是指如果对任何给出的量 ϵ,不管多小,总存在一个正的量 η_0,使得对所有小于 η_0 的正值 η, $f(x \pm \eta) - f(x)$ 都小于 ϵ. 无论人们给于 x 什么样的值,只需 x 和 $x \pm \eta$ 属于从 a 到 b 的区间,这个 η_0 一定对所要求的(性质)同样有效."[28] 然后海涅继续证明了在一个闭区间上连续的函数在此区间上为一致连续,在此证明中他隐含地使用了现在被称作的海涅 - 波莱尔定理. 他还给出了在闭区间上的连续函数达到其最大的最小值这个定理的证明,这是所有证明中最早发表的.



图 16.5　俄国邮票上的索菲亚 · 柯瓦列芙斯卡娅.

魏尔斯特拉斯本人很少公开发表他的许多新思想,正是通过他的追随者和他的学生,其中包括了第一个取得数学博士学位的妇女索菲亚 · 柯瓦列芙斯卡娅(1850—1891) 的努力,才使得魏尔斯特拉斯的许多概念成为数学分析中的标准准则,并且至今仍然有其合适的位置.

人物小传

索菲亚·柯瓦列芙斯卡娅(1850—1891)(Sofia Kovalevskaya)

当一些小姑娘醒来看到她们育婴室墙上娇嫩的花朵时,索菲亚·瓦西列芙娜·科尔芬的房间却糊满了米哈依尔·奥斯特洛格拉茨基的微积分课程的讲义.她的父亲是个军官,喜欢过数学,并允许索菲亚跟一个家庭教师学习.她长大后也喜欢数学.但她不能进一步地学下去,因为她是个女人.俄国的大学不能正式地允许妇女入学,而她的家庭也不允许她单独到欧洲的大学去学习.索菲亚解决这个问题的办法是与弗拉吉米尔·柯瓦列夫斯基定了"有利的婚姻",他是一位科学与政治著作的出版商,而且他自己也是一个有抱负的科学家.

有了丈夫,索菲亚便能出国学习数学,先是在海德堡大学,然后到柏林师从魏尔斯特拉斯.尽管魏尔斯特拉斯和其他一些人为她的事给教职员理事会作了强烈的推荐,柏林大学却不像海德堡大学,它拒绝正式接纳妇女.索菲亚以私人身份跟魏尔斯特拉斯学习,她写了许多可供发表的数学论文,其中最具意义的是关于偏微分方程理论方面的.在这之后,她在1874年接受了哥廷根大学颁发的博士学位,那里愿意对不在校的人颁发学位.

回到俄国后,柯瓦列夫斯基夫妇有了充裕的时间生活在一起并在1878年生了一个女儿.几年以后索菲亚又恢复了由于要支持家务和社会活动而暂时搁置一边的数学研究.因她的丈夫在1883年的逝世,她在斯德哥尔摩大学谋得了一个教授席位,这是当代的第一位女教授.很快她便成了欧洲数学界的一名活跃分子,当了瑞典杂志《数学学报》的编委,又在1888年获得了法国科学院的波登奖,以表彰她在绕定点的刚体旋转方面的革命性工作.

对索菲亚·柯瓦列芙斯卡娅来说,作为一位单身母亲的生活是艰难的.如像她在给朋友的一封信中写的"所有这些愚蠢但又不能耽误的日常事务是对我耐心的严峻考验.我开始明白为什么男人总是高度珍视贤慧能干的家庭主妇了.倘若我是个男人,我必定会为自己选一个漂亮娇小的主妇,她能把我从所有这些事务中摆脱出来."[27] 可惜,柯瓦列芙斯卡娅没有时间来完成她的数学承诺.1891年初在到法国和德国的一次旅途中她患了肺炎并在返回瑞典的几日后逝世.在此后的许多年来,她的墓地一直是旅游的朝觐地,不仅对数学家而言是这样,而且对女权运动的支持者们也是如此(图16.5).

16.2 分析的算术化

甚至在有了波尔查诺和柯西的关于收敛性和连续性的新定义后,在19世纪中叶的一些结果的证明中仍明显缺少了一个至关重要的步骤,这些结果包括了诸如介值定理和有界递增序列的极限存在定理.虽然新定义使数学家们能够证明某个数项序列满足柯西准则,但却没有办法在没有事先指出此极限为什么类型的"数"之前就断定极限的存在性.像其他一些人那样,柯西直观地理解了什么是实数.他甚至曾断言过一个无理数可以看作是某个有理数序列的极限.然而他因此断言这样一种数的存在性是先验的,没有论断来表明这个断言的合法性.

16.2.1 戴德金分割

在世纪中期,许多数学家在积极地考虑准确地说究竟什么是无理数这个问题.他们不再满足于像他们18世纪的先辈做过的那样来假定这种对象的存在性,特别是因为类似的"显然"假定曾导致了不正确的结论.例如,在柏林教授分析基础的过程中,魏尔斯特拉斯构造了一个函数,它处处连续却处处不可微,这是个在上世纪所不能相信有的函数.因此,魏尔斯特拉斯和戴德金开始仔细地思考无理数含意的问题.正如戴德金在他的1858年写就却在1872年才发表的简短的作品《连续性和

无理数》(Stetigkeit und irrationale Zablen) 中的前言所写的:

> 作为苏黎世工科学校教授,我第一次发现我自己不得不讲述微分计算的原理,而且比以前任何时候都强烈感觉到缺乏一个真正的算术的科学基础.在讲解一个变量趋向于一个固定极限值的观念中 …… 我求助于几何的证据.甚至到现在,这种在微分运算的开篇表述里诉诸于几何直观的办法,在我看来还是十分有用的,如果你不想耗费太多时间的话,这也确实是必不可少的.但是这种引进微分运算的形式不能说是科学的,这没有人会否认…….这种叙述是如此经常地让微分运算去处理连续量却没有在任何地方给出这种连续性的解释:甚至在最严谨的微分运算的文章中 …… 所依赖的定理从没有建立在一种纯算术的形式上…….那么只有继续在算术原理中去发现(这些定理的)真正的起源,因而同时也保证了连续性实质的真正定义.[29]

戴德金的研究纲领导致了现今被称作的分析的算术化,它对分析中定理的推导首先来自定义整数的基本公设,然后来自集合论的原理.

为确保他的"连续性的实质" 的定义是"纯算术的方式",戴德金以考虑有理数集 R 的有序性作为开始.他考虑最重要的是,首先如果 $a > b$ 且 $b > c$,则 $a > c$;第二,如果 $a \neq c$,则在 a 与 c 之间存在无穷多个有理数;第三,任意数 a 把整个集合 R 分成两类 A_1 和 A_2,其中 A_1 由小于 a 的数组成而 A_2 由大于 a 的数组成,数 a 本身被指定属于这两个类中的一个.这两个类具有一个显然的性质,即 A_1 中的每个数小于 A_2 中的每个数.为了作出有理数和直线上点的对应关系,戴德金注意到甚至希腊人也清楚的"直线 L 在单个点上较之于有理数领域 R 在单个数方面更为无限地丰富."但是人们不能用几何的直线来算术地定义数.因此戴德金的目标是"创造新的数,使得数的领域像直线那样具有同样的完备性,或者说 … 相同的连续性."[30]

为了创造出新的数,戴德金决定把他所考虑的直线的连续性实质的性质转移到数的领域上来,这个性质即"如果直线的所有点分成了两类,使得第一个类中的每点都在第二个类中每个点的左边,则存在一个且惟一一个点,它产生出这个划分."[31] 因此戴德金把有理数的第三个性质推广到了一个新定义:"如果给出系统 R 的任何一个分成两类 A_1 和 A_2 的划分,它仅具有这个特征性质即 A_1 的每个数 a_1 小于 A_2 中每个数 a_2,则 … 我们称这样一个划分为一个**分割**并以 (A_1, A_2) 表示它."[32] 每个有理数 a 决定了一个分割,对它而言 a 或是 A_1 中数的最大者或是 A_2 中数的最小者,但确实有些分割它不是由有理数产生的,例如定义 A_2 为所有其平方大于 2 的有理数集合,A_1 为所有其他有理数构成的集合而决定的分割."不是所有的分割都是由有理数产生的这个性质在于所有有理数组成的领域的不完备性或不连续性.那么任何时候当我们有一个非有理数产生的分割 (A_1, A_2) 的时候,我们便创造了一个新的数,一个无理数 α,我们把它看为由分割 (A_1, A_2) 完全定义的;我们将说数 α 对应于这个分割或是它生成了这个分割."[33] 戴德金于是把 α 看成是对应于分割概念的一个新创造的数.但其他的数学家觉得最好还是定义这个实数 α 就是这个分割.

无论如何,所有这些分割的集合确定了实数系 $\mathcal{R}$.戴德金能够证明这个系统具有一个自然的序 $<$,它满足有理数所满足的序的同样性质.然后他证明了系统 $\mathcal{R}$ 也具有连续性的属性,即如果它被分裂成两类 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 使得 $\mathcal{A}_1$ 中每个数 α_1 都小于 $\mathcal{A}_2$ 中每个数 α_2,则存在一个且惟一的一个数 α,它或是 $\mathcal{A}_1$ 中最大的数或是 $\mathcal{A}_2$ 中最小的数.事实上,α 是对应于分割 (A_1, A_2) 的实数,其中 A_1 由 $\mathcal{A}_1$ 中所有的有理数组成,而 A_2 由 $\mathcal{A}_2$ 中的所有有理数组成.证明 α 具有所需要的性质是直截了当的事.

戴德金以定义他的新系统 $\mathcal{R}$ 中的标准算术运算来结束了他的文章.因此他有能力来"到达定理的真正证明(像例如,$\sqrt{2}\sqrt{3} = \sqrt{6}$),在我所知的最大范围内它从来没有被建立过."[34] 这些定理之

一是,每个有界递增的实数序列$\{\beta_i\}$有一个极限.设$\mathcal{A}_2$为所有满足对所有i有$\beta_i < \gamma$的数γ的集合,$\mathcal{A}_1$则为剩下的数的集合.戴德金轻易地便证明了分割$(\mathcal{A}_1,\mathcal{A}_2)$确定了数$\alpha$,它是$\mathcal{A}_2$中的最小数,即所要求的极限.

16.2.2 康托尔和基本序列

戴德金关于分割的工作以印刷形式出现于1872年.十分有趣的是,算术地定义实数的问题在那时如此流行,至少有其他四项工作完成了同一目标并在几乎在相同的时间发表,所有这些都基于以满足柯西准则的序列来定义无理数这一普遍的思想.查理·梅瑞(1835—1911)是第一个发表的人(1869),随后在1872年的是艾恩斯特·柯萨克(1839—1902),他给出了关于魏尔斯特拉斯方法的报告.还有乔治·康托尔(1845—1918),以及爱德华·海涅,后者用康托尔的方法推导出了实数的算术基本定理.我们马上要在这里讨论康托尔的方法,因为它对集合论有着深远的含意.

康托尔对待创造实数的问题所依的观点与戴德金不同.他感兴趣的是傅里叶级数收敛性这个老问题,提出的问题是,是否代表一个函数的三角函数级数一定是惟一的.1870年他在三角函数级数在x的每个值上收敛的条件下设法证明了惟一性.然后他成功地减弱了这些条件.首先,在1871年证明了,如果收敛性或这个表达式在所给区间上只有有限个点上不成立时,定理仍然为真.其次,在第二年他已能证明,如果这些例外点的数目无限,但这些点却以一种特别的方式分布,则惟一性仍然成立.为了准确地描述点的这种分布,康托尔意识到他需要一个描述实数的新办法.

像戴德金一样,他也从有理数集开始;康托尔引进了**基本序列**的观念.这是一个序列$a_1,a_2,\cdots,a_n,\cdots$,它的性质是"对任意正有理数$\varepsilon$,存在一个整数$n_1$,使得对$n \geqslant n_1$和任意正整数$m$有$|a_{m+n}-a_n| < \varepsilon$."[35] 现在被称作柯西序列的这个序列满足了柯西在1821年所设定的准则.对柯西而言,这样一个序列收敛于一个实数b是显然的.另一方面,康托尔认识到这样说是犯了一个逻辑错误,因为这样陈述已事先假定了这样一个实数的存在性.因此康托尔用这个基本序列去**定义**一个实数b.换句话说,使每个有理数的基本序列伴随一个实数.有理数r自己也伴随了一个序列,即序列$r,r,\cdots,r,\cdots$,但也有没有伴随有理数的序列.举例说,序列1,1.4,1.41,1.414,…,它是由所熟悉的计算$\sqrt{2}$的算法产生的,就是那样的基本序列.

<table>
<tr><td rowspan="2">人物小传</td><td>乔治·康托尔(1845—1918)(Georg Cantor)</td></tr>
<tr><td>乔治·康托尔有可能跟随着他母亲的音乐家庭的脚步成为一个小提琴家,有时他自己也觉得奇怪为什么竟没有这样.替而代之的是,在圣彼得堡的中小学时期他对数学发生了兴趣,而后从1862年起开始在苏黎世大学学数学,一年以后又在柏林师从魏尔斯特拉斯.只用了十年时间他便成了哈雷大学的正教授,但他渴望在柏林得到一个更高报酬和更受尊敬的位置.反对康托尔的无限集合理论的克罗内克却设法阻止他来首都.在后来的年代里,尽管有精神疾患,他还是能在1890年组织了德国数学会并主要负责了1897在苏黎世召开的第一届国际数学家大会的筹备工作.</td></tr>
</table>

意识到两个基本序列可能都收敛于同一个极限,康托尔便继续定义了在这些序列的集合中的一个等价关系.那么,伴随序列$\{a_i\}$的数b称为等于伴随序列$\{a_i'\}$的数b'是指,如果对任意$\varepsilon > 0$,存在一个n_1,对$n > n_1$成立$|a_n - a'_n| < \varepsilon$.因此,实数的集合$B$便是基本序列等价类的集合.不难在这些序列间定义有序关系,也不难建立基本的算术运算.但是康托尔需要证明的是他所定义的集

合在某种意义上与数直线是一样的.康托尔清楚知道直线上的每个点对应于一个基本序列,但他意识到相反的方向则需要一个公理,即对每个实数(基本序列的等价类)对应于直线上一个确定的点.

在定义了实数之后康托尔回到了他原来的三角级数论的问题上.利用将实数与直线上点的等同性,他定义了一个点集 P 的**极限点**为"位于直线上的那样一个点,在它的每个邻域中我们可以找到无限多个 P 的点 ….一个点的邻域应理解为包含此点在其内部的每个区间.由此容易证明,一个由无限多个点组成的(有界)点集至少有一个极限点."[36] 康托尔以 P' 表示这些极限点的集合,称其为 P 的**导出集**.相似地,如果 P' 是无限的,康托尔定义了第二导出集为 P'' 的极限点集.(如果 P' 有限,则其极限点集为空集.)以这种方式继续下去,康托尔定义了任意有限阶的导出集.于是他区分了两种类型的有界点集.第一类是那些点集,它们对某个 n 值有导出集 $P^{(n)}$ 为空集,而第二类是那些不满足这个条件的集合.例如在$[0,1]$区间中的点集$\{1,1/2,1/3,\cdots\}$的导出集为$\{0\}$,因而为第一类,而此区间中的有理数集以整个区间为导出集因而为第二类.康托尔能够证明当对任意 n,$P^{(n)}$ 为有限集的集合是第一类的从而便能证明对应于一个函数的三角级数在收敛性或表达式不成立的点集为第一类时是惟一的.

16.2.3　无限集

导出集的概念把康托尔引进了一个全新的领地.因为他认识到,对任何集合 P 成立 $P' \supseteq P'' \supseteq P''' \supseteq \cdots$,由此他定义了一个新的集合 Q,即所有 $P^{(n)}$ 的交集.在康托尔原始定义的意义下一般说来 Q 不是个导出集.但是由于它是从 P"导出的",康托尔把它记为 $Q = P^{(\infty)}$.而后他又开始了这个过程,并定义 $P^{(\infty+1)}$ 为$(P^{(\infty)})'$,$P^{(\infty+2)} = (P^{(\infty)})''$,…,甚至于

$$P^{(\infty\cdot 2)} = \bigcap_n P^{(\infty+n)}.$$

因此康托尔引向了考虑当今所称的超限序数,那是些"超越"了有限的序数.

在处理这些点集时康托尔还有另一个疑问.他知道有理数是稠密的但却不是连续的.因而似乎在某种意义上比较起有理数来应该有"更多"的实数.1873 年 11 月,他在给戴德金的信中提出了这个问题:"取所有正整数 n 的集合,记为(n);再设想所有实数 x 的集合,记为(x);问题仅仅是,是否(n)和(x)可以如此对应,使一个集合的每个个体对应于另一个集的一个且惟一的一个个体?就我所知而言,我倾向的意见是(n)和(x)没有这样的独特对应.但我找不出理由来."[37]

戴德金不能回答康托尔的问题,但仅在一个月之后康托尔自己已经能够证明这样的对应是不可能有的.他的证明使用了矛盾法.如果区间(a,b)中的实数可以一一对应于自然数,则能将这些实数排列为 $r_1, r_2, r_3, \cdots, r_n, \cdots$.然后康托尔去找寻区间中的一个实数而它不在这些排列中.他先从此排列中取出头两个数 a', b',满足 $a' < b'$,然后再从(a',b')中取出它的头两个数 a'', b'',满足 $a'' < b''$.按此方式继续进行,他确定出了一个区间套序列(a,b),(a',b'),(a'',b''),….有两种可能性:第一种,这些区间的个数是有限的.这时,在最小的区间$(a^{(n)}, b^{(n)})$中确实有实数,它不在原来的排列中;第二种,如果这些区间的个数为无限,则它们给出了两个有界的单调序列$\{a^{(i)}\}$和$\{b^{(i)}\}$,它们分别有极限 $\bar{a}, \bar{b}$.如果 $\bar{a} \neq \bar{b}$,则区间$(\bar{a}, \bar{b})$中必含有不在原来排列中的实数.最后,如果 $\bar{a} = \bar{b} = \eta$,则 η 也不在原来的排列中.因为,如果对某个 k 它等于 r_k,它就不可能出现在超过了某个指标以后的那些区间中;但由它的定义,作为一个极限它必定包含在所有这些区间中.

在那篇包括了上述证明在内的 1874 年的论文中,康托尔还放进了另一个证明,即所有代数数的集合可以与自然数集建立一一对应关系.由此推出,存在有无穷多个超越数.然而更加重要的是,康托尔建立了一个对无限集合计数的技术,确定了以实数连续统为一方的和以有理数集或代数数

集为另一方的在大小(基数)上的清晰差异.不久之后,在给戴德金的另一封信中他问道,是否可在正方形与区间之间找到一个一一对应.这时的答案显然是"否",事实上康托尔的有些同事觉得这个提问是滑稽可笑的.但在三年后康托尔发现答案应该是"是".他构造了一个对应,将以无限小数展开 $x = a_1 a_2 a_3 \cdots, y = b_1 b_2 b_3 \cdots$,表示的点偶$(x, y)$映到点 z,它可用展开 $z = a_1 b_1 a_2 b_2 a_3 b_3 \cdots$ 来表示.在这个表示中有个小小的问题,它与 0.19999… 和 0.2000… 表示了相同的数这个事实有关,但康托尔很快纠正了他的证明,建立了这个一一对应关系的存在性.虽然这个结果令数学界感到惊奇,但戴德金指出康托尔的映射是不连续的,因而它与维数的几何意义没有什么关系.事实上,不久许多数学家给出了不可能有这种连续的一一映射存在的证明,因此维数是在连续的一一对应下的不变量.

16.2.4　集合论

不久康托尔便认识到,他的一一对应的概念可以被置于一个新集合论的基础之中.在 1879 年他用它开始了无限集合的基数的研究,这是个最终联系到超限序思想的一个概念.定义两个集合 A 和 B 为有**相同势**为:在 A 的元素与 B 的元素之间存在一个一一对应.最初康托尔举出两个特殊的情形,即与自然数集 N 有相同势的集合,称这些为可数集,以及那些与实数集具有相同势的集合.对连续统的性质作进一步了解的努力导致了康托尔在随后二十多年中建立了无限集的详细的理论.这个理论的许多内容都发表在 1895 年和 1897 的两篇论文中,它们总合在一起,冠以《对超限集合论创建的贡献》(Beiträge zur Begründung der transfiniten Mengenlehre)的标题.

《贡献》以集合的一个定义,即"一个把我们具体的或想像中确定而又各自分离的对象 m 收集为一个整体的 M."[38]"每个集合 M 有一个确定的'势',我们也称它为它的'基数'",他接着说"我们要以 M 的'势'或'基数'的名字来称呼这个普遍概念,借助于我们活跃的思考功能,当我们对它的各种元素性质以及所赋与它们的序作出抽象时,这个概念便从集合 M 中产生了."[39] 康托尔尔此处所说的"抽象"的意思是,一个无限集的基数是对有限集"元素个数"概念的推广.因此自然数集合和实数集合具有不同的基数.自然数集的基数是具有最小超限基数的集合,康托尔把此基数称作"阿列夫零",记为 $\aleph_0$,而实数集的基数记为 C.如果两集合间存在一个一一对应,则称它们等价或具有相同基数.康托尔也在超限基数间定义了 $<$ 的观念.集合 M 的基数$\bar{\bar{M}}$ 小于 N 的基数是指,如果没有 M 的任何部分等价于 N 而有 N 的一个部分等价于 M.很清楚,对两个集合 M 和 N 而言最多只有 $\bar{\bar{M}} = \bar{\bar{N}}, \bar{\bar{M}} < \bar{\bar{N}}$ 或 $\bar{\bar{N}} < \bar{\bar{M}}$ 中一个关系能成立.但康托尔并没有指出至少有其中一个关系必须成立.

由于 $\aleph_0 < C$,康托尔提出了问题:是否对实数的子集具有其他基数的可能.在 1878 年,他以为他已对此问题得到了否定的答案:"通过归纳过程,我们对此不作进一步描述,人们可以得到如下定理,即直线的点集间由这个(等价)产生的类的个数有限,实际上它等于二."[40] 实数的每个点集具有基数 $\aleph_0$ 或 C 的这个猜想被称为**连续统假设**.虽然康托尔多次相信他有了对此结论的证明,而且至少有一次相信他有了一个否定性的证明,但是他和其他任何人都没有能够证明或反向证明这个假设.事实上,连续统假设最终被指出,用集合论的任何合理的公理系统都是不可证明的.

尽管康托尔不能回答所有他提出来的与无限集有关的问题,他的对这些集合的概念既得到了广泛的承认也得到了强烈的批评.特别是列奥波德·克罗内克,他相信任何的数学构造必须要能够在有限次数学运算中完成才行.由于康托尔的一些数学构造不满足克罗内克的准则,作为克雷尔的《杂志》的一个编委,克罗内克把康托尔的一篇文章拖延了很长的时间,以致于康托尔拒绝再在《杂志》上发表文章了,即便如此,它仍是那时最有影响的数学杂志.虽然克罗内克和其他一些人继续反对康托尔的超限方法,仍然有不断增多的数学家们支持他的处理集合论的方式.这两群人之间的

冲突一直延续到了今天.

16.2.5　戴德金和自然数的公理

康托尔已经发展出了一些集合论中更加先进的思想,而且与戴德金一起指出了如何从有理数来构造实数.但正是戴德金借助于集合来描述自然数从而有理数的特征的办法完成了分析算术化的进程.在他成就了这项任务的著作中,他也给出了对集合论基本观念的一个导论.这里所说的任务是弄清什么是(自然)数,它们是什么意思?这花去了他超过 15 年的时间,直到 1888 年才发表.

为了描绘自然数的特性,戴德金从这样的观念出发,即自然数构成一个**事物**的或是"我们想像的对象"的集合.因此,戴德金定义了**系统**这个词,这里翻译为**集合**:"对不同的事物 $a,b,c,\cdots$ 以某种理由可以用一种共同的观点来考虑,这是常常发生的事;这些不同的事物便能在思想中被联系起来,我们就说它们形成了一个集合 $S\cdots$.这样的一个集合 S 作为我们想像中的一个对象就像一件东西一样;当就每件东西而言它被确定了是否是 S 的一个元素时,这个集合 S 就被完全确定了."[41] 在给出了这个必定有点含糊的定义后,戴德金描述了涉及集合的各种简单关系.例如,一个集合 A 是集合 S 的一**部分**是说 A 中每个元素也是 S 的一个元素.还有,由任意集合 $A,B,C,\cdots$ **合成**的集合是由那些至少在集合 $A,B,C,\cdots$ 中一个里的元素组成的,记为 $M(A,B,C,\cdots)$;而 $A,B,C,\cdots$ 的公共元素组成的集合被记为 $G(A,B,C,\cdots)$.在现代的术语中,戴德金的"部分"变成了我们的"子集",$M(A,B,C,\cdots)$ 变成了集合 $A,B,C,\cdots$ 的"并"而 $G(A,B,C,\cdots)$ 是他们的"交".

自然数的一个基本性质是如下的观念,即对每个数存在一个惟一的后继元.换句话说,有一个从自然数集 N 到自身的函数 ψ,满足 $\psi(n)=n+1$.一般地,戴德金定义 S 上的一个函数为一个"规则,对 S 中每个确定的元素 s,按照它有属于它的一个确定的事物,称它为 S 的**变换**,并记为 $\varphi(s)$."[42] 因为 N 的不同元素有不同的后继元,戴德金由此引出了**相似**或**内射**变换的观念,即一个这样的变换:"对 S 中不同的元素 a,b 总对应于不同的变换 $a'=\varphi(a)$, $b'=\varphi(b)$"[43].在这种情形下,存在系统 $S'=\varphi(s)$ 的逆变 $\bar{\varphi}$,定义为对每个 S' 中元素 s' 派定为那个由 φ 变换成它的惟一的元素 s.两个集合 R 和 S 被称为相互**相似**是指如果存在一个定义在 R 上的内射变换 φ 使得 $S=\varphi(R)$.

自然数另外的性质是,在后继元变换下 N 的像是 N 自己的一个真子集,惟一不属于这个像的是元素 1.事实上,集合 N 的无限性恰寓于这个像为真子集的性质中:"当一个集 S 相似于它自己的一个真子集时称它为无限的,相反的情形时则说 S 是有限集".[44] 但是到底存在无限集吗?在没有论证这类集合的存在性前戴德金不太愿意去证明有关这些集合的结果.于是他给出了一个例子:"所有能成为我想像的对象的东西全体 S 是无限的,因为如果 s 表示为 S 中元素的意思,那么想像 s',即 s 能成为我想像的对象,它自己也是 S 的一个元素"[45].给出从 S 到自身的变换,定义为 $s\to s'$,对此,其像不是整个 S 是清楚的,并且这变换是内射的.戴德金得到结论是集合 S 确实满足他的定义的要求.

戴德金认识到,N 具有一个内射的后继元函数,其像为 N 的真子集,但这些性质并不能惟一地描绘出 N 的特性.很可能在任意集合 S 中有额外的元素满足这些性质,而这些元素并非自然数.例如,正有理数集满足所有这些性质.故而戴德金又加上了另一个性质,即一个元素属于 N 当且仅当它是 S 的每个那样的子集 K 的元素,其中 K 满足如下性质:1 属于 K 且 K 中每个元素的后继元也属于 K.换句话说,N 是由它是其所有满足原来性质的所有集合的交作为特征刻画出来的.因此 N 包含了一个基本元素 1,1 的后继元 $\varphi(1)$,以及它的后继元 $\varphi(\varphi(1))$,等等,而没有其它的元素.

由他对自然数的特征描绘,戴德金能够推导出的除了数学归纳原理外,还给出一个定义并导出 N 上的有序关系,还有加法和乘法运算.两个另外的数学家,吉乌塞甫 · 佩亚诺(1858—1932) 和哥

特洛伯·弗雷格(1848—1925)也在1880年代考虑了构造自然数并导出其重要性质这些相同的问题.弗雷格的工作以印刷形式出现于1884年,而佩亚诺的工作则发表于1889年.正是这个工作和魏尔斯特拉斯及其学派的工作一起,使得微积分学能被置于一个从集合论基本观念出发的坚实基础之上.因此表明了微积分的存在是独立于运动及曲线的物理世界的,而这个世界正是牛顿首先用来创立了这门学科.

16.3　复分析

回想1837年,哈密顿发展出了复数理论,把复数看作是一对有序的实数,从而给出了神秘的-1的平方根究竟是什么这个问题的一个答案.虽然自16世纪以来数学家们一直在使用复数,但是即便在哈密顿的工作之后一般地还不能以这种抽象的形式来构想它们.1797年挪威的税务官卡斯帕·韦塞尔(1745—1818)在一篇文章中首先发表了这些数的几何表示法,它最终成了考虑复数量的新方法的基础,这个方法很快使数学家们相信,他们可以放心地使用这些数了.

16.3.1　复数的几何表示

韦塞尔在他的《关于方向的分析表示》中的目的最初并不是要与复数相关联.他觉得如果有一种办法能以单独的代数表达式来表示平面中的一个线段的长度和方向,那么某些几何的概念会被更加清楚地理解.韦塞尔明白这些表达式必须要能够被代数地处理.特别地,他需要一个方法来代数地表达出方向的一个任意变化,而且要比简单地使用负号来表示相反方向广泛得多.

韦塞尔从处理加法开始:"两段直线相加是说,如果我们以这种方式把它们连接起来,即第二段直线开始于第一段的终端,然后从被连接的线的第一点作一条直线到它的终点.这条线段便是被连接的线段的和."[46] 因此,不管线段的代数表达是什么,两条线段的加法必须满足由韦塞尔的运动概念中提取出来的这个显然的性质.换句话说,他设想线段代表了向量.然而正是乘法给了韦塞尔对他的方向表示问题的基本答案.为了推导出乘法,他建立了许多个他觉得是本质的性质.首先平面中两线段的积必须仍在此平面中.其次,乘积线的长度必须是两个因子长度的积.最后,如果所有的方向都以正向单位直线量起,则乘积的方向角应是这两个因子方向角的和;他以1表示那条正向单位线.以ε记垂直于直线1的单位线段,他容易地证明了他所希望的性质意味着$\varepsilon^2=(-\varepsilon)^2=-1$或$\varepsilon=\sqrt{-1}$.一条具单位长度并与正单位线夹角为$\theta$的线段现在可以表示成$\cos\theta+\varepsilon\sin\theta$,而且一般地,长度$A$,夹角$\theta$的线可以表示为$A(\cos\theta+\varepsilon\sin\theta)=a+\varepsilon b$,其中$a$和$b$为适当选取的线段(图16.6).因此复数的几何解释是从韦塞尔的几何线段的代数解释中产生的.加法的显见的代数规则满足于韦塞尔对那些运算的要求,而乘法$(a+\varepsilon b)(c+\varepsilon d)=ac-bd+\varepsilon(ad+bc)$满足他的乘法公理.韦塞尔也轻易地从他的定义中推导出复数乘法和开根的标准规则.

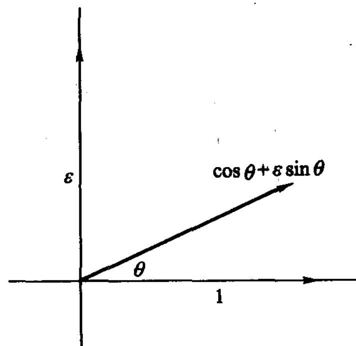


图16.6　韦塞尔的复数的几何表示.

可惜韦塞尔的文章在发表后的许多年内没有在欧洲的大部分地区被读到.同样的命运也等待着瑞士的书商让·罗伯特·阿尔冈(1768—1822),他在1806年出版的一本小书中提出了复数的类似几何解释.只是由于高斯在证明代数基本定理中和在研究

四次剩余中(图 16.7)使用了与他们相同的复数的几何解释后,这种解释才赢得了数学界的承认.高斯对基本定理即实系数多项式 $p(x)$ 有一个实或复根如此着迷,以至他发表过对它的四个证明,它们分别发表于 1799,1815,1816 和 1848 年.每个证明都在某种形式上使用了复数的几何解释,尽管在头三个证明中以分别考虑复数的实、虚部的方式隐藏了这个观念.这样,在他最初的证明中,高斯实际上令 $p(x+\mathrm{i}y)=u(x,y)+\mathrm{i}v(x,y)$,并留意到 p 的一个根就是曲线 $u=0$ 和 $v=0$ 的交点.因此他详细研究了这些曲线,并通过运用介值定理证明了这些曲线必定相交.只是在他 1848 年的最后一个证明中,高斯相信数学家们会对复数的几何解释有了相当舒服的感觉,故而他可以明确地使用它.事实上,在那个证明中,类似于第一个证明那样,他甚至允许多项式的系数为复数.

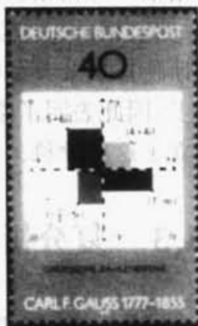


图 16.7　德国邮票上的高斯复平面.

16.3.2　复积分

在本世纪的第二个十年中,高斯以他对复数意义的清晰理解开始发展复变函数理论.在 1811 年给他朋友弗里德里希·魏尔海姆·贝塞尔(1784—1846)的一封信中,高斯不仅讨论了复数的几何解释而且也讨论了 $\int_u^v \varphi(x)\mathrm{d}x$ 的意义,其中变量 x 是复的:

> 我们必须假定 x 从积分等于 0 处到 $x=a+b\mathrm{i}$ 通过了许多无穷小增量(每个具有 $\alpha+\beta\mathrm{i}$ 的形式),然后将所有这些 $\varphi(x)\mathrm{d}x$ 加起来.这样做的意思是完全确定的.但此过程却有无穷多的不同方式;这恰如人们把实数量的整个区域想成一条无限的直线一样,从而可以使得所有量的区域,包括实的和虚的量,含有无限平面的意思,其中每个点由代表量 $a+b\mathrm{i}$ 的横坐标 a 及纵坐标 b 决定.由 x 的一个值到另一个值 $a+b\mathrm{i}$ 的这个连续过程便依此沿一条曲线进行,从而可能有无穷多种方式.[47]

高斯进而给出了那个"非常漂亮的定理",即只要在连接积分起点和终点的两条不同曲线所包围的区域中 $\varphi(x)$ 不是无穷的,则此积分沿此两条曲线的值便相同.虽然高斯没有用那些术语,但他正是把 $\varphi(x)$ 当成了一个解析函数.不管怎样,他从来没有发表过这一结果的证明.但此结果的证明却由柯西于 1825 年发表,因此此定理一般被称作柯西积分定理.

柯西首先在 1814 年的一篇论文中考虑了复区域上的积分问题,但该论文却到 1827 年才发表.在这篇著作中他主要感兴趣的是计算定积分的值,这些积分的限或是一端或是两端为无穷.为完成这个计算,他力图使用欧拉和拉普拉斯发展起来的各种严格的方法,以便把积分路径移动到复平面中.特别地用欧拉的思想导出了柯西-黎曼方程.在大约 1777 年的一篇文章中欧拉断言,对复变函数最重要的定理是对每个写成和 $M(x,y)+\mathrm{i}N(x,y)$ 的函数 $Z(x+\mathrm{i}y)$ 具有性质 $Z(x-\mathrm{i}y)=M-\mathrm{i}N$.这时可推出,如果

$$V=\int Z\mathrm{d}z=\int(M+\mathrm{i}N)(\mathrm{d}x+\mathrm{i}\,\mathrm{d}y)$$
$$=\int M\mathrm{d}x-N\mathrm{d}y+\mathrm{i}\int N\mathrm{d}x+M\mathrm{d}y=P+\mathrm{i}Q,$$

那么,以 $x-\mathrm{i}y$ 替换 $x+\mathrm{i}y$,则有

$$P-\mathrm{i}Q=\int(M-\mathrm{i}N)(\mathrm{d}x-\mathrm{i}\,\mathrm{d}y)$$

$$= \int M\mathrm{d}x - N\mathrm{d}y - \mathrm{i}\int N\mathrm{d}x + M\mathrm{d}y.$$

因此,$P = \int M\mathrm{d}x - N\mathrm{d}y$,$Q = \int N\mathrm{d}x + M\mathrm{d}y$,其中积分号代表反微分,这是欧拉常用的符号.因为 P 是微分 $M\mathrm{d}x - N\mathrm{d}y$ 的积分,从而

$$\frac{\partial M}{\partial y} = -\frac{\partial N}{\partial x}.$$

同样地,对 Q 的表达式得出

$$\frac{\partial M}{\partial x} = \frac{\partial N}{\partial y}.$$

这两个被称作柯西 - 黎曼方程的方程最终成了复变函数的特征性质.

在他 1821 年的书《分析教程》中,像欧拉做过的那样,柯西以分别考虑实部和虚部的办法来处理复数量.因此他考虑了"符号表达式"$a + \mathrm{i}b$,并以正常的代数规则把它们乘在一起就"好像$\sqrt{-1}$是个其平方为 - 1 的实数量一样."[48] 他借助两个实变量的两个实函数定义一个复变函数,并指出复区域上的各种标准的超越函数是什么意思.然后他将他的关于级数收敛性的大部分结果推广到了复数,其中用复量 $z = a + \mathrm{i}b$ 的模$\sqrt{a^2 + b^2}$作为实数绝对值的类比.他还用两个构成函数的连续性来定义复变函数的连续性.

但是,一直到了 1825 年,在发现了他的定积分的新定义之后,柯西才有可能从它们自身出发来处理复变函数.在他的《论取虚界限的定积分》(Mémoire sur les intégrales définies prises entre des limites imaginaires)中,他明晰地定义了复定积分

$$\int_{a+\mathrm{i}b}^{c+\mathrm{i}d} f(z)\mathrm{d}z$$

为"形如$[(x_1 - a) + \mathrm{i}(y_1 - b)]f(a + \mathrm{i}b)$,$[(x_2 - x_1) + \mathrm{i}(y_2 - y_1)]f(x_1 + \mathrm{i}y_1)$,$\cdots$,$[(c - x_{n-1}) + \mathrm{i}(d - y_{n-1})]f(x_{n-1} + \mathrm{i}y_{n-1})$的乘积的和,当两个序列 $a, x_1, x_2, \cdots, x_{n-1}, c$ 和 $b, y_1, y_2, \cdots, y_{n-1}, d$ 由从第一个到最后一个递增或递减的项构成并在它们的项数无限增加而彼此无限地趋近时,它们收敛的极限或极限之一."[49] 换句话说,柯西直接地取两个区间$[a, b]$和$[c, d]$的划分来推广他对函数定积分的定义.但是像高斯一样,柯西也认识到始于 $a + \mathrm{i}b$ 终于$c + \mathrm{i}d$ 的不同积分路径有无数条.因此不清楚这个定义是否有意义.为了证明他的积分定理,实际上也就是说他的定义是有意义的.他从考虑由参数方程 $x = \varphi(t)$,$y = \psi(t)$ 定义的路径开始,其中 φ 和 ψ 为 t 在区间$[\alpha, \beta]$中的单调可微函数,而 $\varphi(\alpha) = a$,$\varphi(\beta) = c$,$\psi(\alpha) = b$,$\psi(\beta) = d$.那么两个序列$\{x_j\}$和$\{y_j\}$由单个序列 $\alpha, t_1, t_2, \cdots, t_{n-1}, \beta$ 确定,只要各自用 φ 和 ψ 来计算在它们的值即可.假定由 t_j 确定的各种子区间的长度是小的,柯西注意到成立

$$x_j - x_{j-1} \approx (t_j - t_{j-1})\varphi'(t_j),$$

及

$$y_j - y_{j-1} \approx (t_j - t_{j-1})\psi'(t_j).$$

从而这个定积分便是形如

$$(t_j - t_{j-1})[\varphi'(t_j) + \mathrm{i}\psi'(t_j)]f[\varphi(t_j) + \mathrm{i}\psi(t_j)]$$

的这些项和的极限,因此可以写成如下形状:

$$\int_{a+\mathrm{i}b}^{c+\mathrm{i}d} f(z)\mathrm{d}z = \int_{\alpha}^{\beta}[\varphi'(t) + \mathrm{i}\psi'(t)]f[\varphi(t) + \mathrm{i}\psi(t)]\mathrm{d}t,$$

或者,令 $x' = \varphi'(t)$,$y' = \psi'(t)$ 时写成

$$\int_{\alpha}^{\beta}(x' + \mathrm{i}y')f(x + \mathrm{i}y)\mathrm{d}t.$$

“现在假定只要 x 在 a 与 c 的界限之间而 y 在 b 与 d 界限之间，函数 $f(x+\mathrm{i}y)$ 保持有界和连续. 在个情况下容易证明积分的值…与函数 $x=\varphi(t), y=\psi(t)$ 的性质无关.”[50] 柯西对此陈述的证明建立在变分计算的基础之上，证明要求 $f'(z)$ 的存在性与连续性，而柯西还不曾明晰定义过一个复变函数的导数是什么意思. 柯西以 $\varphi+\varepsilon u, \psi+\varepsilon v$ 替代 φ 和 ψ 来对曲线作无穷小变动，其中 ε 是“一个一阶无穷小量”，u 和 v 在 $t=\alpha$ 和 $t=\beta$ 时为零，并把在积分中的相应变化展开为 ε 的幂级数. 利用分部积分法，柯西证明在此级数中 ε 的系数为 0，从而积分路径的一个无穷小变动产生了积分的一个 ε^2 级的无穷小变化. 柯西得出了结论说，路径的有限变动，即从一条积分路径变到第二条这种路径，在积分中产生的不过是无穷小变化，就是说，根本没有变化. 对柯西来说，积分定理因而得证，虽然不是现代标准的证明.

下一步柯西考虑 f 在某个值 $z_1=r+\mathrm{i}s$ 变为无穷的情形，这里 $a\leqslant x\leqslant c, b\leqslant y\leqslant d$. 沿着两条连接起来包围了 z_1 的路径上的积分不再是相同的. 定义 R 为 $\lim_{z\to z_1}(z-z_1)f(z)$，柯西计算出沿着两条彼此无限接近，同时也无限接近点 z_1 的路径积分之差为 $2\pi R\mathrm{i}$. 例如，如果 $f(x)=1/(1+z^2)$，则 f 在 $z=\mathrm{i}$ 时为无穷. 因为

$$\lim_{z\to\mathrm{i}}\frac{z-\mathrm{i}}{1+z^2}=\lim_{z\to\mathrm{i}}\frac{z-\mathrm{i}}{(z+\mathrm{i})(z-\mathrm{i})}=\frac{1}{2\mathrm{i}},$$

从而这个函数在沿从 -2 到 2 的两条图 16.8 所示的路径 L_1 和 L_2 的积分之差是

$$2\pi\frac{1}{2\mathrm{i}}\mathrm{i}=\pi.$$

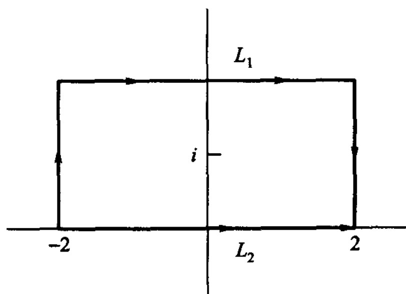


图 16.8　对 $f(z)=\dfrac{1}{1+z^2}$ 的两条从 -2 到 2 的积分路径.

在 1826 年的一篇文章中，柯西在某种程度上推广了他的积分定理. 给出 $f(z)$ 为无穷大的值 z_1，柯西看出 $f(z_1+\varepsilon)$ 对 ε 幂的展开式以负幂次开始. 展式中 $1/\varepsilon$ 的系数被柯西称为 $f(z)$ 在 z_1 的**留数**，记以 $R(f,z_1)$，于是，如果 $(z-z_1)f(z)=g(z)$ 在 z_1 附近有界，则

$$f(z_1+\varepsilon)=\frac{g(z_1+\varepsilon)}{\varepsilon}=\frac{1}{\varepsilon}g(z_1)+g'(z_1+\theta\varepsilon),$$

θ 为 0 与 1 之间的一个数. 由此得到，$f(z)$ 在 z_1 的留数为 $g(z_1)$，即前面以 R 表示的那个值.

柯西知道，他的留数理论对一些问题有所应用，像有理分式的分解，确定某些定积分的值，还有某些类方程的解. 举例来说，他证明了

$$\int_{-\infty}^{\infty}\frac{\cos x}{1+x^2}\mathrm{d}x=\pi\,\mathrm{e}^{-1},$$

用的办法是把积分区间扩大为复平面上包含使被积函数为无穷的 i 值的一条闭曲线. 计算的中心思想是，在由一个半圆和实轴上一个区间围成的路径上的积分可由留数算出，而当半圆的半径(以

及区间的长度)变大时,在半圆上的那部分积分趋向于0.

16.3.3 复变函数和线积分

复变函数论中有许多其他的标准结果至少应该部分地归功于柯西,其中的大多数是他的积分定理和留数计算的应用.我们将对他1846年的一篇文章作简短的分析来作为对他的工作讨论的结束;虽然那篇文章根本没有提到复变函数,但却引出了证明积分定理的一个新办法并提供了在向量分析和拓扑学中某些基础思想的起端.这篇短文《关于可以拓展到一条闭曲线上所有点的积分》(Sur les intégrales qui s'étendent a tous les points d'une courbe fermée)包含了许多光有陈述而无证明的定理.柯西承诺以后将给出证明.显然他并没有这样做.这些定理处理了一个多变量 $x,y,z,\cdots$ 的函数 k,对它沿着一个曲面 S 的边缘曲线 Γ 积分,而 S 在一个没有指明维数的空间中.最重要的一些结果被收集在下面的定理中.

定理 假设

$$k = X\frac{\mathrm{d}x}{\mathrm{d}s} + Y\frac{\mathrm{d}y}{\mathrm{d}s} + Z\frac{\mathrm{d}z}{\mathrm{d}s} + \cdots,$$

其中 $X\mathrm{d}x + Y\mathrm{d}y + Z\mathrm{d}z + \cdots$ 是一个恰当微分.(一个微分是恰当的指 $\partial X/\partial y = \partial Y/\partial x, \partial X/\partial z = \partial Z/\partial x, \partial Y/\partial z = \partial Z/\partial y, \cdots$.)假设函数 k 在 S 上除有限多个内部的点 $P, P', P'', \cdots$ 外处处有限和连续.如果 $\alpha, \beta, \gamma, \cdots$ 为 S 中各自包围这些点的闭曲线,则

$$\int_\Gamma k\mathrm{d}s = \int_\alpha k\mathrm{d}s + \int_\beta k\mathrm{d}s + \int_\gamma k\mathrm{d}s + \cdots.$$

特别地,如果没有这些奇点,则

$$\int_\Gamma k\mathrm{d}s = 0.$$

在二维的情形,S 是平面的一个区域,k 为一任意微分,则

$$\int_\Gamma k\mathrm{d}s = \pm\iint_S\left(\frac{\partial X}{\partial y} - \frac{\partial Y}{\partial x}\right)\mathrm{d}x\mathrm{d}y.$$

如果 k 为恰当微分,则 $\partial X/\partial y = \partial Y/\partial x$,故其右端为零,从而左端也为零.

柯西积分定理由最后的断言得到.一个复变函数 $f(z) = f(x + \mathrm{i}y)$ 可以表示为 $f(x,y) = u(x,y) + \mathrm{i}v(x,y)$,并且由于 $\mathrm{d}z = \mathrm{d}x + \mathrm{i}\,\mathrm{d}y$,有

$$\int f(z)\mathrm{d}z = \int(u\mathrm{d}x - v\mathrm{d}y) + i\int(v\mathrm{d}x + u\mathrm{d}y).$$

柯西-黎曼方程表明这两个被积项为恰当微分,因此积分定理成立.

但是比积分定理更有意思的是在柯西的文章中出现了 n 维空间中的线积分的概念(实际上是对维数大于3的空间中的)以及定理中的那个今天广为人知的格林定理的论述(倒数第二句).事实上,多少类似于这个定理的结果出现在乔治·格林(1793—1841)1828年的一篇论文中,它处理的是电磁学的问题,但是柯西的文本是现今教科书上被如此称呼的那个结果的第一个出版了的叙述文字.最后,把绕曲面边缘的线积分表示为绕孤立奇点的线积分之和,标志了研究积分与曲面之间关系的开端,其中积分在此曲面上并非处处有定义.我们称那些绕孤立奇点的积分值为**周期**.由于柯西从没有发表过他1846年的定理的证明,人们只能对于他能把所有他的新概念推进到多远作些推测.但是,在几年之后,是黎曼重述了柯西的结果,并给予完全的证明,还将关于周期的结果扩大到远远超过了柯西的概念.

16.3.4　黎曼和复变函数

黎曼的论文《单复变函数一般理论的基础》(Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse)一开始便讨论了实变和复变函数之间的一个重要区别.虽说函数的定义"对每个(变量 z 的)值对应于未定量 w 的一个单值",[51] 可同时用于实和复的情形,但黎曼认识到,在后一种情形即 $z = x + \mathrm{i}y, w = u + \mathrm{i}v$,定义导数的比值 $\mathrm{d}w/\mathrm{d}z$ 的极限很大地依赖于 $\mathrm{d}z$ 是怎样趋向于0的.由于对从代数方式定义的函数人们可以形式地来计算导数,并不会发生此类问题,故而黎曼决定把导数的存在性当作复函数概念的一个基础性东西:"当复变量 z 的变化使得导数 $\mathrm{d}w/\mathrm{d}z$ 的值与 $\mathrm{d}z$ 的值无关时,称复变量 w 为另一个复变量 z 的函数."[52] 当然,柯西在他对复变函数的整个讨论中也使用了这个观念,但只是在他生涯的最后阶段才把它明确起来.

人物小传	乔治·伯恩哈德·黎曼(1826—1866)(Georg Bernhard Riemann)
	在1846年当他被哥廷根大学录取后,黎曼需要在他父亲的允许下才从学习神学和哲学转到了学习数学.他的生活是从布雷策伦兹村开始的,这里距汉堡东南50英里,现在他要到柏林去,因为哥廷根的数学教育不是特别强.在柏林他遇见了狄利克雷,并以他为导师.几年之后他返回了哥廷根师从高斯学习并于1851年得到了博士学位.他进行了两年的研究工作并为他在哥廷根大学取得教学资格的演讲作准备.1857年他被任命为副教授,两年之后,由于其间也来到哥廷根大学的狄利克雷的逝世,他被任命为正教授.他的数学工作是光辉的,然而肺结核病缩短了他的工作,并于1866年在他为寻求治疗到意大利的一次旅程中失去了他的生命.

作为这个定义的第一个应用,黎曼证明了当把这样的复变函数考虑成从 z - 平面到 w - 平面的一个映射时它是保角的.假设 p' 和 p'' 在 z - 平面中无限靠近原点 P,则它们的像 q' 和 q'' 无限靠近 P 的像 Q.记 p' 到 P 的无穷小距离为 $\mathrm{d}x' + \mathrm{i}\,\mathrm{d}y'$,也记为 $\varepsilon' e^{i\varphi'}$,而 q' 到 Q 的距离记为 $\mathrm{d}u' + \mathrm{i}\,\mathrm{d}v'$ 或 $\eta' e^{i\psi'}$,以相似的记号表示其他的无穷小距离,黎曼注意到他对函数赋予的条件表明

$$\frac{\mathrm{d}u' + \mathrm{i}\,\mathrm{d}v'}{\mathrm{d}x' + \mathrm{i}\,\mathrm{d}y'} = \frac{\mathrm{d}u'' + \mathrm{i}\,\mathrm{d}v''}{\mathrm{d}x'' + \mathrm{i}\,\mathrm{d}y''},$$

或者是

$$\frac{\mathrm{d}u' + \mathrm{i}\,\mathrm{d}v'}{\mathrm{d}u'' + \mathrm{i}\,\mathrm{d}v''} = \frac{\eta'}{\eta''} e^{i(\psi' - \psi'')} = \frac{\mathrm{d}x' + \mathrm{i}\,\mathrm{d}y'}{\mathrm{d}x'' + \mathrm{i}\,\mathrm{d}y''} = \frac{\varepsilon'}{\varepsilon''} e^{i(\varphi' - \varphi'')}.$$

由此推导出 $\eta'/\eta'' = \varepsilon'/\varepsilon''$ 和 $\psi' - \psi'' = \varphi' - \varphi''$,或者说无限小三角形 $p'Pp''$ 与 $q'Qq''$ 相似.这样一个保持角度的映射被叫作一个**保形**映射.在某种意义上说,欧拉和高斯已知道解析复函数有此性质,但只是黎曼给出了这个论证,而且另外他还能证明黎曼映射定理,它是说复平面上任意两个单连通区域可以用一个适当选取的复函数把它们相互地共形映射到对方.

然后黎曼推导出了柯西 - 黎曼方程,所用的方法是借助于两个函数 u 和 v 来确定出导数的存在性所意味着的东西:

$$\frac{\mathrm{d}w}{\mathrm{d}z} = \frac{\mathrm{d}u + \mathrm{i}\mathrm{d}v}{\mathrm{d}x + \mathrm{i}\mathrm{d}y} = \frac{\frac{\partial u}{\partial x}\mathrm{d}x + \frac{\partial u}{\partial y}\mathrm{d}y + \mathrm{i}\left(\frac{\partial v}{\partial x}\mathrm{d}x + \frac{\partial v}{\partial y}\mathrm{d}y\right)}{\mathrm{d}x + \mathrm{i}\mathrm{d}y}$$

$$= \frac{\left(\frac{\partial u}{\partial x} + \mathrm{i}\frac{\partial v}{\partial x}\right)\mathrm{d}x + \left(\frac{\partial v}{\partial y} - \mathrm{i}\frac{\partial u}{\partial y}\right)\mathrm{i}\,\mathrm{d}y}{\mathrm{d}x + \mathrm{i}\mathrm{d}y}.$$

如果这个值与 $\mathrm{d}z$ 如何趋向于 0 无关,那么依次令 $\mathrm{d}x$ 和 $\mathrm{d}y$ 为 0 并将所得到的表达式的实部和虚部相等便证明了

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y} \quad 及 \quad \frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y}.$$

反之,如果这些柯西-黎曼方程得到满足,则可容易算出所求导数为 $\partial u/\partial x+\mathrm{i}\partial v/\partial x$,它是与 $\mathrm{d}z$ 无关的值.黎曼把这些方程和由它们容易得到的第二组偏微分方程

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \quad 及 \quad \frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}=0$$

一起作为复变函数论的中心.

作为例子,黎曼依照柯西在 1846 年所提出的概述给出了对柯西积分定理的一个详细证明.重要的思想是格林定理,黎曼把它叙述为下面的形式:

定理 设 X 和 Y 为 x,y 的两个函数,在有限区域 T 中连续,以 $\mathrm{d}T$ 表示其无穷小面积元.则

$$\int_{\mathrm{T}}\left(\frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}\right)\mathrm{d}T=-\int_{\mathrm{S}}(X\cos\xi+Y\cos\eta)\mathrm{d}s,$$

其中后一个积分取在 T 的边界曲线 S 上,ξ,η 代表曲线的内法线分别与 x 轴和 y 轴的夹角.

黎曼证明这个定理的办法是使用微积分的基本定理,沿着平行于 x 轴的直线对 $\partial X/\partial x$ 积分,在这些直线穿过区域边界的地方得到 X 的值.由于在每个这些点上 $\mathrm{d}y=\cos\xi\mathrm{d}s$,他便能对 y 积分,得到了

$$\int\left[\int\frac{\partial X}{\partial x}\mathrm{d}x\right]\mathrm{d}y=-\int X\mathrm{d}y=-\int X\cos\xi\,\mathrm{d}s.$$

定理另外的一半按相似方法证明.然后,黎曼注意到

$$\frac{\mathrm{d}x}{\mathrm{d}s}=\pm\cos\eta \quad 及 \quad \frac{\mathrm{d}y}{\mathrm{d}s}=\mp\cos\xi,$$

这里的符号取决于从切线到内法线是逆时针还是顺时针进行的.由此推出,格林定理可重新写成

$$\int_T\left(\frac{\partial X}{\partial x}+\frac{\partial Y}{\partial y}\right)\mathrm{d}T=\int_S\left(X\frac{\mathrm{d}y}{\mathrm{d}s}-Y\frac{\mathrm{d}x}{\mathrm{d}s}\right)\mathrm{d}s,$$

由此便容易得到柯西积分定理.

黎曼论文的大部分内容涉及到复变函数研究中一个全新的概念,即黎曼面的思想.在实的单变量函数情形,可以用二维空间中一条曲线来画出这个函数.对复变函数这种表示不再能行,因为图像需要在实的四维空间中.图示复变函数的另一种方法是在一个平面中让自变量 z 沿一条曲线变动并考虑由非独立变量 w 在另一平面中所生成的曲线.黎曼从一个复变函数总有一个幂级数表示这个事实认识到"一个定义在(x,y)平面上一个区域中的 $x+\mathrm{i}y$ 的函数能以惟一的方式被解析地开拓."从而推出,一旦人们知道了在某个区域中的值,便可开拓此函数,甚至于,譬如说沿一条返回到同一 z 值的连续曲线.于是有两种可能性,"取决于被开拓函数的性质,或是不管怎样开拓函数对相同的[z]值总是取相同的值,或者不是这种情形."[53] 在第一种情形,黎曼称此函数为单值的,而第二种情形,称其为多值的.对后者的一个简单例子是取 $w=z^{1/2}$.为有效地研究这类函数,不可能只简单用上述的两个平面,因为人们不知道对第一个平面上的已知点函数会取哪一个值.因此黎曼想了一个新的主意,即使用一个多重平面,它是 z 平面的一个多层的覆盖,而函数就在上面取值.这些层沿一条直线连接在一起,譬如说沿负实轴,使得一旦在一条曲线中移动并穿过那条直线时,则函数值就从一层变到另外一层.这样一来多值函数在此黎曼面的每个点上只确定了一个值.由于可能发生这样的情况,在多次轮转后(上述例子中是两次)又回到了原先的值,那么覆盖的最上一层

必须与最底下一层连接起来.由此可知,一般说来不可能在三维空间中构造一个黎曼面的具体模型.不过,由黎曼处理多值复函数作为开端的对黎曼面的研究立刻把黎曼和其他一些人引进了一个今天被称为**拓扑学**的领域.1846 年的柯西所完全没有接触到的拓扑与沿曲线和曲面积分之间的关联在 19 世纪后半期和 20 世纪初期得到了极其详尽的探索.

16.4　向量分析

黎曼在 1851 年叙述的格林定理是用了一个二重积分和一个沿曲线的并对曲线的长度元 ds 积分的相等性.在物理中曲线上积分的运用代表了沿此曲线所做的功,它似乎促使了在符号使用上的变化;这发生在 1850 年代,这时曲线积分被形如$\int pdx + qdy$ 的线积分所取代.虽然这个符号已经用于复积分,但物理学家却把它转换为涉及向量的表达式.在 19 世纪中涉及到向量的其他物理概念引出了其他重要的积分定理.

16.4.1　线积分和多重连通性

克拉克·麦克斯韦在 1855 年注意到,如果 α、β 和 γ 各为"电作用强度"ϵ 平行于 x -, y - 和 z - 轴的分量,而 l, m, n 为此曲线的切线对应的方向余弦(即切线与三个坐标轴间夹角的余弦),则 ϵ(考虑为沿曲线作用)可以写为形式 $l\alpha + m\beta + n\gamma$.因为 $lds = dx$, $mds = dy$ 及 $nds = dz$,麦克斯韦写作$\int \epsilon ds = \int \alpha dx + \beta dy + \gamma dz$.第二年,这个符号出现在查利斯·德劳内(1816—1872)的物理教科书中.德劳内比麦克斯韦多少更明确一些,他注意到,如果 F 是力而 F_1 是其沿一条曲线的切分量,则沿此曲线作用的力所做的功可用$\int F_1 ds$ 表示.另外,如果 F 的直角分量为 X, Y, Z,则后一积分可重写为$\int Xdx + Ydy + Zdz$.

线积分的符号很快成为物理学中一个标准符号,并被黎曼在 1857 年的一篇文章中采用,文中研究了(黎曼)曲面 R,在上面画出了一些被线积分所取的曲线.黎曼一开始就看到了当在曲面上的一个区域的周界上对恰当微分 $Xdx + Ydy$ 取积分时,它为零:

> 在两个固定点之间取两条不同的路径,并假定这两条路径合起来围成了 R 中一个区域的整条边界,那么沿这两条路径所取的积分$\int (Xdx + Ydy)$ 有相同的值.因此,如果在 R 内部的每条闭曲线都围成了 R 中一个区域,那么当从一个起点到一个相同的终点取积分的话,则此积分总是同一个值,并且是终点位置的连续函数,它与积分路径无关.这给出了曲面中的一个特别的类:单连通曲面,在它上面每条闭曲线都界定了曲面的一个区域,例如一个圆盘;而多重连通曲面上则不会发生上述情况,它的例子是由两个同心圆所围成的圆环.[54]

黎曼把多重连通性的观念进行了精炼:"一个曲面 F 被称为$(n + 1)$重连通的是指在它上面可画上 n 条闭曲线 $A_1, \cdots, A_n$,它们中没有一个单独或组合起来能够围成 F 的一个区域,但如果增加任一个其他的闭曲线 A_{n+1} 时,这一组闭曲线围成了 F 的某个区域."[55]黎曼进一步注意到,一个$(n + 1)$重连通的曲面经过一次切割后可以变成一个 n 重连通曲面,而这里的切口曲线是从一个边缘点出发穿过其内部然后到达另一个边缘点.举例说,一个圆环是二重连通的,它可以经由任何一个不

断开它的切割把它变成单连通区域.而一个双重圆环则需要两次切割才能变成单连通区域.

利用切割的想法,黎曼能够准确地描述出当人们在一个$(n+1)$重连通曲面R上对一个恰当微分进行积分时所发生的事.如果人们从这个曲面上取下几条切割曲线,便剩下了一个单连通曲面R'.从一个固定点出发的R'的任意曲线上对恰当微分$X\mathrm{d}x+Y\mathrm{d}y$积分则像上面一样,确定了一个曲面上位置的单值连续函数Z.但是每当积分路径跨过一条切口时,它的值便跳跃了一个固定的数,这个数依赖于这个切口,一共有n个这样的数,每一个对应一个切口.多重连通的观念被证明在物理学中是重要的,特别在流体力学和电磁学中,故而它被一些物理学家推广到了三维空间.这些物理学家包括了海尔曼·冯·赫姆霍兹(1821—1894),威廉·汤姆逊(1824—1907)和麦克斯韦.

16.4.2 曲面积分和发散量定理

物理学家们不仅对线积分有兴趣而且对于曲面积分,函数积分以及二维区域上的向量场感兴趣.回忆一下,早在1760年拉格朗日就已经在计算曲面面积时给出了曲面面积元$\mathrm{d}S$的显式表达.但是一直到了1811年在他的《分析力学》的第二版中拉格朗日才引进了曲面积分的一般观念.他注意到,如果在$\mathrm{d}S$的切平面与xy-平面间的夹角为γ,则简单的三角计算可将$\mathrm{d}x\mathrm{d}y$重写为$\cos\gamma\mathrm{d}S$.从而如果A是个三元函数,则$\int A\mathrm{d}x\mathrm{d}y=\int A\cos\gamma\mathrm{d}S$,其中第二个积分是取在曲面的一个区域上的.类似地,如果β是这个切平面与xz-平面做成的角而α是与yz-平面做成的角,则$\mathrm{d}x\mathrm{d}z=\cos\beta\mathrm{d}S$及$\mathrm{d}y\mathrm{d}z=\cos\alpha\mathrm{d}S$.拉格朗日知道$\alpha,\beta$及$\gamma$也可以看成是曲面的法线与$x$-,$y$-和$z$-轴之间各自所成的夹角.

拉格朗日利用曲面积分来处理流体力学.高斯在1813年利用同样的概念来考虑一个椭球体的重力吸引力.但是高斯比拉格朗日走得更远,他指出了在曲面S由三个方程$x=x(p,q),y=y(p,q),z=z(p,q)$参数地给出时应该如何计算对$\mathrm{d}S$的积分.利用一个几何的推断,他证明了

$$\mathrm{d}S=\left[\left(\frac{\partial(y,z)}{\partial(p,q)}\right)^2+\left(\frac{\partial(z,x)}{\partial(p,q)}\right)^2+\left(\frac{\partial(x,y)}{\partial(p,q)}\right)^2\right]^{1/2}\mathrm{d}p\mathrm{d}q,$$

并由此知道,任何一个对$\mathrm{d}S$的积分可以化成形如$\int f\mathrm{d}p\mathrm{d}q$的积分,其中$f$是两个变量$p,q$的函数,或是显式的或是隐式的.

高斯运用他对曲面上积分的研究证明了一个特殊情形下的结果,即今天所知的散度定理.这个定理的一般情形是由米哈依尔·奥斯特罗格拉茨基(1801—1861)在1826年叙述并证明的;奥斯特罗格拉茨基是位俄国数学家,1820年代他正在巴黎学习.[56] 在一篇出自他对热学理论研究的标题为《积分学中一个定理的证明》的文章中,奥斯特罗格拉茨基考虑了一个曲面,它围成了具体积元ω的立体区域,而自身的面积元为ε.以p,q和r为x,y,z的三个可微函数,而角α,β,γ如上述那样定义,奥斯特罗格拉茨基把散度定理叙述为下面的形式:

$$\int\left(\frac{\partial p}{\partial x}+\frac{\partial q}{\partial y}+\frac{\partial r}{\partial z}\right)\omega=\int(p\cos\alpha+q\cos\beta+r\cos\gamma)\varepsilon,$$

其中左边的积分取在立体V上,而右边的取在其边缘曲面S上.在现今,这个定理依拉格朗日的主意,一般地写成形式

$$\iiint_V\left(\frac{\partial p}{\partial x}+\frac{\partial q}{\partial y}+\frac{\partial r}{\partial z}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iint_S p\mathrm{d}y\mathrm{d}z+q\mathrm{d}z\mathrm{d}x+r\mathrm{d}x\mathrm{d}y.$$



图 16.9 俄国邮票上的奥斯特罗格拉茨基.

人物小传	米哈依尔·奥斯特罗格拉茨基(1801—1861)(Mikhail Ostrogradsky) 米哈依尔·奥斯特罗格拉茨基(图 16.9)是通过想成为一名军官的愿望而找到了他进入数学之路的.由于他生于乌克兰的一个普通的下层人家,不能承受那种没有独立收入的军官生活方式的昂贵花费.为了他未来的生活他在 1816 年进入了哈尔科夫大学.他对数学和物理变得感兴趣,并在 1820 年通过了学位考试.但实际上他并未得到学位,因为宗教事务和国家教育大臣决定处罚奥斯特罗格拉茨基的老师 T.F.奥谢波夫斯基,他是大学的神学院长.理由是他不能对他的学生灌输正宗的宗教和拥戴沙皇的看法. 奥斯特罗格拉茨基离开俄国到巴黎去了数年,在那里他写出了一些他的最重要的数学著作.1828 年他回到了圣彼得堡并被选为科学院院士.他到一些军事学院教数学以酬他原先的军事方面的雄心壮志.1847 年在这些学校里负责所有的数学教育,后来写了供在那里使用的许多重要的教科书.

像格林定理一样,这个结果是微积分基本定理的一个推广,故而奥斯特罗格拉茨基的证明使用了那个定理.为了在一个以 x- 方向穿过该立体并具横截面积$\bar{\omega}$ 的"狭长圆柱"上对$(\partial p/\partial x)\omega$ 积分,他用基本定理表达该积分为$\int (p_1 - p_0)\bar{\omega}$,其中 p_0 和 p_1 为 p 在圆柱与立体相交的曲面的一片上的值(图 16.10).因为在曲面的一个截面上 $\bar{\omega} = \varepsilon_1 \cos \alpha_1$,而在另一截面上为 $\bar{\omega} = -\varepsilon_0 \cos \alpha_0$,其中 α_1 和 α_0 各为在曲面元 ε_1 和 ε_0 的法线所构成的角,奥斯特罗格拉茨基便证明了

$$\int \frac{\partial p}{\partial x}\omega = \int p_1\varepsilon_1 \cos \alpha_1 + \int p_0\varepsilon_0 \cos \alpha_0 = \int (p \cos \alpha)\varepsilon,$$

其中左边的积分是取在圆柱上而右端的取在曲面的两片上.把在所有这些圆柱上的积分加起来便

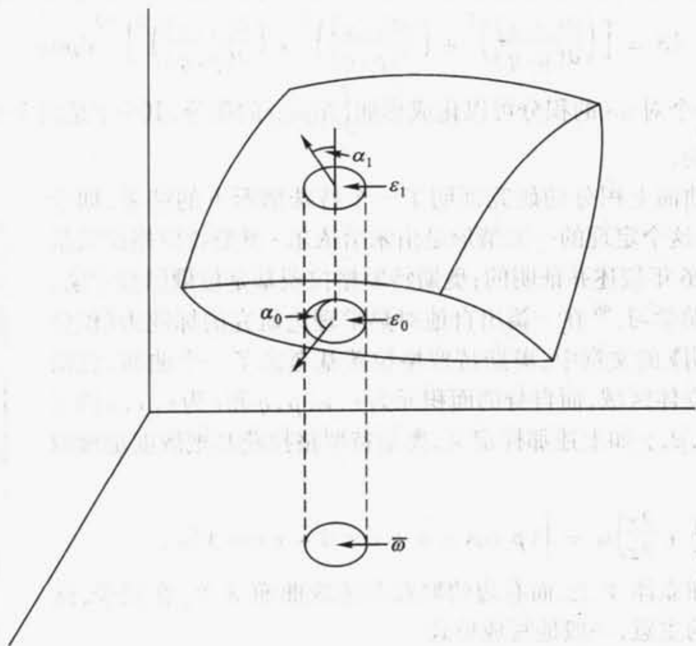


图 16.10　奥斯特罗格拉茨基对散度定理的证明.

给出了所要结果的三分之一部分,另外三分之二的部分以相似方法得到.十分有意思的是,奥斯特罗格拉茨基在 1836 年把他的结果推广到了 n 维的情形,从而在维数大于 3 的情形中给出了对这个结果的最早的叙述.

16.4.3 斯托克斯定理

散度定理把立体上的积分与在边缘曲面上的积分联系了起来,而格林定理则把平面中一个区域上的积分与在边缘曲线上积分联系起来.一个类似的结果是对三维空间中一个曲面上的积分和绕边界曲线上的积分进行比较,这个现在被称作斯托克斯定理的结果首先以印刷形式出现于 1854 年.乔治·斯托克斯(1819—1903)许多年来一直在剑桥大学设立了史密斯奖竞赛,在 1854 年 2 月的竞赛中问了下面的问题:

问题 8 如果 X,Y,Z 为直角坐标 x,y,z 的函数;dS 是任一有限曲面的元素;l,m,n 是 dS 的法线与这些坐标轴倾角的余弦;ds 为边缘曲线的元素,证明

$$\iint\left[l\left(\frac{\partial Z}{\partial y}-\frac{\partial Y}{\partial z}\right)+m\left(\frac{\partial X}{\partial z}-\frac{\partial Z}{\partial x}\right)+n\left(\frac{\partial Y}{\partial x}-\frac{\partial X}{\partial y}\right)\right]dS$$
$$=\int\left(X\frac{dx}{ds}+Y\frac{dy}{ds}+Z\frac{dz}{ds}\right)dS,$$

…… 单重积分取于曲面的整个边缘上.

不知道是否有哪个学生证明了这个定理,但麦克斯韦的确坐在考场中.然而这定理就已经出现在威廉·汤姆逊给斯托克斯的一封 1850 年 7 月 2 日的信中,而左边的那个被积项也已出现在斯托克斯早先的两部著作中,在那里它表示了某个流体的角速度.这个结果的证明首次发表在 1861 年海尔曼·汉克尔(1839—1873)的一篇论文中,至少是在曲面被 $z=z(x,y)$ 的显式表达情形下如此.汉克尔把 z 的值和 $dz=(\partial z/\partial x)dx+(\partial z/\partial y)dy$ 代入右端的积分,从而把它化为一个双变量的积分,然后运用格林定理把它转换成一个二重积分,这时容易看出它等于左端的曲面积分.

人物小传

乔治·斯托克斯(1819—1903)(George Stokes)

虽然他的三个兄弟随他的父亲在他们的故乡爱尔兰从事了传教士的职业,然而斯托克斯却在他的一位老师影响下被吸引到了数学上.1837 年他进入了剑桥大学的彭布鲁克学院,在那里他的数学教育大多数都来自他的导师威廉·霍普金斯.斯托克斯于 1841 年以数学荣誉考试第一名的资格毕业,八年之后被任命为卢卡斯数学教授席位,逝世前他一直在这个职位上.在他的职业生涯中他的理论的和实验的研究中涵盖了自然哲学的许多方面,其中包括的领域有流体动力学,弹性和光绕射.他对纯数学的各种涉猎是出于他需要发展出解决特别的物理问题的方法或者要证实他已用过的数学工具的有效性.斯托克斯在各种官方岗位上服务于科学界.特别地,从 1854 年到 1885 年他是皇家学会的秘书,从 1885 年到 1890 年为它的主席,另外从 1887 年到 1891 年还是剑桥大学在议会中的代表.

斯托克斯证明了一个相关的结果:显然有

$$\frac{\partial}{\partial x}\left(\frac{\partial Z}{\partial y}-\frac{\partial Y}{\partial z}\right)+\frac{\partial}{\partial y}\left(\frac{\partial X}{\partial z}-\frac{\partial Z}{\partial x}\right)+\frac{\partial}{\partial z}\left(\frac{\partial Y}{\partial x}-\frac{\partial X}{\partial y}\right)=0.$$

斯托克斯在 1849 年证明了等同于上述等式的逆结果,即如果 A,B,C 为满足

$$\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}=0$$

的函数,则存在函数 X, Y,和 Z 使得

$$A = \frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}, \quad B = \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}, \quad C = \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}.$$

像克莱罗在二维情形证明的结果即一个恰当微分是一个函数的微分那样,这个结果只对某些简单的区域成立.汤姆逊在 1851 年也给出了此结果的不同证明,但不管是斯托克斯还是汤姆逊都没有处理对区域的限制条件.他们的证明需要某个微分方程是可解的,但他们只简单地假定这个解是可以解出的而不费心确保这一点的特定条件.无论怎样,他们的结果与斯托克斯定理本身一起表明,在所述的关于 A, B 和 C 的条件下,曲面积分 $\iint (lA + mB + nC)\mathrm{d}S$(或用更现代的符号为 $\iint A\mathrm{d}y\mathrm{d}z + B\mathrm{d}z\mathrm{d}x + C\mathrm{d}x\mathrm{d}y$)不依赖于曲面而只依赖于边界曲线.

斯托克斯定理和散度定理二者都出现在麦克斯韦的《电磁通论》(Treatise on Electricity and Magnetism)的开篇一章中,并在后面的章节中经常使用它们.由于麦克斯韦是在物理学中使用四元数符号的提倡者,故而他将这些定理写成四元数的形式,使用的事实是,如果向量算子 $\nabla = (\partial/\partial x)\mathrm{i} + (\partial/\partial y)\mathrm{j} + (\partial/\partial z)\mathrm{k}$ 作用于向量 $\sigma = X\mathrm{i} + Y\mathrm{j} + Z\mathrm{k}$,得出的四元数可写成

$$\nabla\sigma = -\left(\frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z}\right) + \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}\right)\mathrm{i} + \left(\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}\right)\mathrm{j} + \left(\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right)\mathrm{k}.$$

出于他对它们的物理意义的解释,麦克斯韦把 $\nabla\sigma$ 的标量部分和向量部分分别命名为 σ 的**收敛量**和 σ 的**旋量**.麦克斯韦的收敛量是今天被称为 σ 的**散度**的负量.

这些定理的纯向量形式最终在接近世纪末时出现在吉布斯的著作中.令 $\mathrm{d}V = \mathrm{d}x\mathrm{d}y\mathrm{d}z$ 为体积元,$\mathrm{d}a = \mathrm{d}y\mathrm{d}z\mathrm{i} + \mathrm{d}z\mathrm{d}x\mathrm{j} + \mathrm{d}x\mathrm{d}y\mathrm{k}$ 为曲面面积元,及 $\mathrm{d}\boldsymbol{r} = \mathrm{i}\,\mathrm{d}x + \mathrm{j}\,\mathrm{d}y + \mathrm{k}\,\mathrm{d}z$,吉布斯把散度定理写成形式

$$\iiint \nabla \cdot \sigma \mathrm{d}V = \iint \sigma \cdot \mathrm{d}\boldsymbol{a}$$

而斯托克斯定理写成

$$\iint (\nabla \times \sigma) \cdot \mathrm{d}\boldsymbol{a} = \int \sigma \cdot \mathrm{d}\boldsymbol{r}.$$

注意,这两个定理中的左端被积项也可分别写为 $\mathbf{div}\,\sigma\mathrm{d}V$ 和 $\mathbf{curl}\,\sigma \cdot \mathrm{d}\boldsymbol{a}$.

即便是向量形式也不能明显看出可以把格林定理,斯托克斯定理和散度定理联合成一个单一的结果.但是维托·沃尔泰拉(1860—1940)却于 1889 年在他研究 n 维空间中的超曲面过程中把它们统一起来了.(在 1836 年对 n 维空间的研究是件新事而在 50 年后已是很普通的了.)沃尔泰拉不仅使用了丰富的指标来陈述一个结果,使得上面提到的三个定理都是其低维的特殊情形,而且他还与亨利·庞加莱(1854—1912)一起把斯托克斯和汤姆逊的被庞加莱称为可积性条件的结果推广到高维的情形,而这些可积性条件是用来保证线积分,曲面积分和它们的高维类比不依赖于它们取积分的曲线,曲面或超曲面而只依赖于那些几何对象的边界.正是这些以庞加莱引理知名的推广为庞加莱研究这些多重积分和积分区域的拓扑之间的关系提供了工具,而这种研究在黎曼那时就已经开始了.庞加莱在接近 20 世纪更替时的一系列文章把这个研究发展成为现在称为代数和微分拓扑学这门学科的开端,它的一些面貌我们将在第 18 章中考察.

16.5　概率论与统计学

19 世纪看到了统计方法在各种领域应用的开端,特别是在农业和社会科学中的应用.事实上

正是这些应用导致了各种标准的统计技巧的发展,它们的大多数发生在20世纪初.在这一节里我们从最早的统计方法之一最小二乘法开始,然后在对拉普拉斯的对概率论和统计学的综合作短暂浏览之后,考虑在19世纪中叶正态曲线的新解释,并考察19世纪最后十来年中统计过程的某些发展.

16.5.1 勒让德和最小二乘法

或许19世纪最重要的统计方法是最小二乘法.这个方法提供了被称作"观测组合"的主要工具之一,它是一种把对一个特殊事件的大量观测度量集合为一个单一的"最佳"结果的方式.例如,假定某人知道一个特定的物理关系被表示为线性函数 $y = a + bx$.他对所考虑的现象进行了多次观测并取得了数据点$(x_1, y_1), (x_2, y_2), \cdots, (x_k, y_k)$.把这些 k 对数依次代入方程中的x 和y 得到了 k 个方程,并用来确定两个未知系数 a 和b.两个未知量的 k 个线性方程的方程组对于解是过多了,一般说来,它没有确定的解.那么,想法是在某程度上确定一个"最佳"的近似解.用几何语言说,这个问题是找一条直线,它在某种意义下以"最接近"的方式通过这 k 个观测点.

观测组合的问题在18世纪被各种各样的数学家讨论过,主要是有关天文观测的.在那些努力解决这个问题的人中有大约在1715年的罗杰·柯茨,1749年的欧拉,1750年的托比亚斯·梅耶尔(1723—1762)和1760年的罗杰·博斯科维奇(1711—1787).柯茨在他身后才发表的一篇文章中简要地描述了他的思想;这种想法是对各种观测给予加权并计算加权平均.欧拉在研究木星和土星之间相互引力作用对各自轨道影响时,最后得到一个八个未知量的75个方程的方程组.他用解出他的方程中各个较小的分组并把这些解组合起来的方法力图找出一个最佳解.另一方面,梅耶尔在观察月亮的细微运动时不得不去解一个三个未知量的27个方程的方程组.他发展出一套系统的办法,把它的方程分为三组,每组九个方程式,同时分别在每组中把方程式加起来而后去解所得到的三个未知量的三个方程式的方程组.不完全清楚的是,按什么确切的准则用来划分这些方程式.是博斯科维奇,当他处理涉及地球真实形状的问题时在这方面做出了有意义的进展.他说出了确定这些方程组的解的方法应该满足的实际准则,其中包括了把一组特定值代入方程式时所产生的误差的绝对值的和进行极小化的重要准则.几年之后,拉普拉斯在他自己的关于同一问题的著作中把博斯科维奇的方法转换成了一个详细的代数方法.

虽然不知道是什么影响了阿德瑞恩·玛瑞·勒让德发展出他的1805年的更好的方法,但正是勒让德的最小二乘法作为解决线性方程的超定方程组的最佳方法流传到现在.这个方法作为附录出现在一本关于确定慧星轨道的著作中.

勒让德以给出引进这个方法的理由作为他的讨论的开端:"所研究的大多数问题都是由观测给出的量度来得到它们能提供的最准确的结果,但这时几乎总产生形如 $E = a + bx + cy + fz + \cdots$ 方程的方程组,其中 $a, b, c, f, \cdots$,是已知的系数,它们从一个方程到另一个方程是有变动的,而 $x, y, z, \cdots$,是未知的,它们必须根据将每个方程 E 化为一个或为0的或很小的量的条件来确定."[58] 用更现代的术语说,即有一个 n 个未知量的m 个方程的系统$(m > n)$, $V_j(\{x_i\}) = a_{j0} + a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = 0 (j = 1, 2, \cdots, m)$,人们需要求出它的"最佳"近似解 $\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n$.对每个方程而言,值 $V_j(\{\bar{x}_i\}) = E_j$ 是对那些解的相伴误差.像他的前辈那样,勒让德的目的是将所有 E_i 变小:"在为此目的所建议的所有原理中,我以为没有比我们在前面研究中所运用的原理更广泛,更准确和更易于应用的了;我们的原理是赋与误差的平方和以极小.由此意味着在这些误差中建立了一种均衡性,它阻止了极端情形所施加的过份的影响;这非常好地适用于揭示出最接近于趋向真实情形的系统

的状态."[59]

为了确定误差平方的最小值,勒让德运用了微积分工具.就是说,为使平方和 $E_1^2+E_2^2+\cdots+E_m^2$ 在 x_1 变动时有一个最小值,则它对 x_1 的偏导数必须为 0:

$$\sum_{j=1}^{m} a_{j1}a_{j0}+x_1\sum_{j=1}^{m} a_{j1}^2+x_2\sum_{j=1}^{m} a_{j1}a_{j2}+\cdots+x_n\sum_{j=1}^{m} a_{j1}a_{jn}=0.$$

因为对 $i=2,3,\cdots,n$ 有类似的方程,勒让德注意到他现在有了对 n 个未知量 x_i 的 n 个方程,因而方程组可以用"现成的方法"解出.虽然他没有由基本原理出发把这个方法推导出来,但他的确看出他的方法是对单个量的一组观测求通常平均值方法的一个推广.在这种情形(即 $n=1$,对所有 $a_{j1}=-1$ 的情形),如果我们令 $b_j=a_{j0}$,则误差的平方和是

$$(b_1-x)^2+(b_2-x)^2+\cdots+(b_m-x)^2.$$

使此和为极小的方程是 $(b_1-x)+(b_2-x)+\cdots+(b_m-x)=0$,故解为

$$x=\frac{b_1+b_2+\cdots+b_m}{m},$$

它正是 m 个观测的普通平均值.

16.5.2　高斯和最小二乘法的推导

在勒让德文章发表的十年间,最小二乘法是整个欧洲大陆在解决天文和大地测量问题的一个标准方法.特别,它出现在高斯的《天体运动理论》(Theoria motus corporum celestium) 中.但是高斯没有引用勒让德的结果.事实上,高斯宣称,自 1795 年以来他自己一直在使用这个原理.高斯的言论引起了勒让德痛心的反应,他提醒说科学发现的优先权只能以出版物确定.他们的争执延续了多年.迟至 1827 年,勒让德还继续严斥高斯剽窃了他人的发现.[60]

不管优先权的争执怎样,高斯较之于勒让德确实把这方法推进得更远.首先他认识到仅仅说人们能用"现成的方法"解出由最小二乘法得到的 n 个变量的 n 个方程的方程组是不够的.在实际应用中,常常有许多个方程而且其系数不是整数而是计算到很多位小数的实数.在这些情形中使用克莱姆法则需要巨大数量的计算.因而高斯设计了对方程组的系统消去方法,这是个把方程乘以适当选取的数然后再把这些新方程加起来的方法.这个过程以高斯消去法为人所知,它实际上等同于1800年前中国汉朝的方法,最后得到的是一个三角形的方程组,就是说它的第一个方程只含有一个未知量,第二个含两个未知量,等等.因此由第一个方程很容易解出它的惟一的未知量,将解代入第二个方程从而解出第二个未知量,等等,直到方程组被完全解出.高斯的方法在世纪后期被德国的大地测量学家威尔海姆·约当(1842—1899)作了一些改进,当时他在处理测量问题中使用了最小二乘法.约当设计了一种替换方法,一旦找到了三角形的方程组,就可以进一步简约为一个对角形的方程组,那里每个方程只含一个未知量.这个高斯 - 约当方法在现代线性代数课程中是被典型地当作解线性方程组的标准方法来教授的.[61]

其次,高斯详尽阐述了最小二乘法为合理的根据,这较之于多少有些含糊的勒让德的"一般原理"要好得多.这就是他由他早先发现的一个适当的函数 $\varphi(x)$ 推导出了这个方法,其中 $\varphi(x)$ 描述了量 x 在确定观测量时误差的概率.此前托马斯·辛普森在 1755 年及拉普拉斯在 1770 年代都曾努力去发现一个合理的误差函数.前者在一篇向皇家学会宣读的文章中试图证明,观测误差会以取多次观测的平均值而减小.为证明这点,例如,他假定在一次特定的天文测量中以秒来度量的误差大小 $-5,-4,-2,-1,0,1,2,3,4,5$ 出现的概率分别与 $1,2,3,4,5,6,5,4,3,2,1$ 成比例.因此单个误

差不超过 1 秒的概率是 16/36 = 0.444,不超过 2 秒的是 24/36 = 0.667.另一方面,他计算出六个误差的平均值不超过 1 秒的概率是 0.725,不超过 2 秒的是 0.967,因而证明了取平均值的好处.

辛普森试图将取平均的结果推广到更一般的误差函数.但他受到了托马斯·贝叶斯的批评,后者评注说"你用一个不完美的工具作的观测越多则越能确定你最终得到的误差将会与你所使用的工具的不完美程度成比例."[62] 然后,拉普拉斯在 1770 年代也试图推导出一个经受得住更仔细分析考验的合理的误差函数,办法是对这种函数应当满足的条件作出明晰的假定.这些条件是,首先,$\varphi(x)$ 应对零是对称的,这是因为在回答贝叶斯的批评中明确地假定了所使用的工具要使得一个观测量过大和过小的可能性是等同的.第二,这曲线必须在两个方向上都渐近于实轴,这是因为一个无限大误差的概率为 0.第三,在 $\varphi(x)$ 下方的总面积应为 1,这是由于在任意两值之间的曲线下方的面积代表了观测具有的误差在这两个值之间的概率.可惜,有许多条曲线都满足拉普拉斯的要求.通过运用其他各种推断,拉普拉斯定下了曲线是 $y = (m/2)e^{-m|x|}$,其中 m 为某个正数.但是拉普拉斯很快就发觉,基于这个误差函数的计算引起了极大的困难.虽然他因此而试过其他的函数,却是高斯在 1809 年成功地对拉普拉斯的问题得出了新的答案.

高斯以与拉普拉斯同样的基本准则出发.他将这些准则加进到勒让德原先的问题中,即确定 n 个未知量 $x_1, x_2, \cdots, x_n$ 的 m 个线性函数 $V_1, V_2, \cdots, V_m$ 的值.假定这些值的观测值为 $M_1, M_2, \cdots, M_m$,其相应的误差为 $\Delta_1, \Delta_2, \cdots, \Delta_m$;高斯知道,由于各种观测是相互独立的,因而所有这些误差出现的概率为 $\Omega = \varphi(\Delta_1)\varphi(\Delta_2)\cdots\varphi(\Delta_m)$ 而找出最有希望的这些值就表明要将 Ω 极大化,但要做到这点需要对 φ 有更好的了解.故而高斯作出更进一步的假设,即"如果在相同的环境和相等的管理下对任一个量经由多次直接观测确定,则这些观测的算术平均值是最希望要的值."[63] 把每个 V_i 取为一个变量的最简单线性函数,即 $V_i = x_1$,他假定观测的算术平均值 $x_1 = (1/m)(M_1 + M_2 + \cdots + M_m)$ 给出了 Ω 的最大值,并由此来确定 φ.当极大值存在时对所有 i 有 $\partial\Omega/\partial x_i = 0$.由于 Ω 是个乘积,高斯把这些方程式换作 $\frac{\partial}{\partial x_i}(\log \Omega) = 0$,或

$$\frac{\frac{\partial}{\partial x_i}(\varphi(\Delta_1))}{\varphi(\Delta_1)} + \frac{\frac{\partial}{\partial x_i}(\varphi(\Delta_2))}{\varphi(\Delta_2)} + \cdots + \frac{\frac{\partial}{\partial x_i}(\varphi(\Delta_m))}{\varphi(\Delta_m)} = 0. \tag{16.5}$$

对每个 j, $\frac{\partial}{\partial x_i}(\varphi(\Delta_j)) = \frac{\partial \varphi}{\partial \Delta_j}\frac{\partial \Delta j}{\partial x_i}$.因为对 $i > 1$ 有 $\frac{\partial \Delta_j}{\partial x_i} = 0$,及对所有 j, $\frac{\partial \Delta_j}{\partial x_1} = -1$,则(16.5)的 n 个方程都约化为单一的方程式

$$\frac{\varphi'(\Delta_1)}{\varphi(\Delta_1)} + \frac{\varphi'(\Delta_2)}{\varphi(\Delta_2)} + \cdots + \frac{\varphi'(\Delta_m)}{\varphi(\Delta_m)} = 0, \tag{16.6}$$

其中以 $\varphi'(\Delta_j)$ 代替了 $\frac{\partial \varphi}{\partial \Delta_j}$.为进一步简化,高斯假定每个观测值 $M_2, M_3, \cdots, M_m$ 都等于 $M_1 - mN$, N 为某个数.得出 $x_1 = M_1 - (m-1)N$, $\Delta_1 = M_1 - x_1 = (m-1)N$,及对 $i > 1$ 有 $\Delta_i = -N$.把这些值代入方程(16.6)便给出了高斯的关系式

$$\frac{\varphi'[(m-1)N]}{\varphi[(m-1)N]} = (1-m)\frac{\varphi'(-N)}{\varphi(-N)}.$$

因为这个关系式对所有正整数 m 均成立,高斯是最后得出

$$\frac{\varphi'(\Delta)}{\Delta\varphi(\Delta)} = k,$$

k 为某个常数,因而 $\log(\varphi(\Delta)) = \frac{1}{2}k\Delta^2 + C$,或者 $\varphi(\Delta) = Ae^{(1/2)k\Delta^2}$.拉普拉斯对 φ 的条件使高斯

断定 k 为负值,让 $k = -2h^2$,最后

$$\varphi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2\Delta^2}.$$

对高斯而言,这就是“正确的”误差函数,因为他可以由此容易地推导出最小二乘法. 毕竟,在给出函数 φ 之后,一般情形下的 Ω 可以由下面表达式给出

$$\Omega = h^m \pi^{-(1/2)m} e^{-h^2(\Delta_1^2+\Delta_2^2+\cdots+\Delta_m^2)}.$$

要使 Ω 极大,因而必须使 $\sum\Delta_i^2$ 极小,就是说,使误差的平方和极小,它正是勒让德发展出来的方法.

误差的分布是“正态”的,就是说是由高斯的函数确定的,这一点由于立刻得到许多按经验得出的证据的支持,变得更加可信了. 特别是弗里德里希 · 贝塞尔对几百颗星球作了三组观测,并比较了按照正态规律在给定范围内的理论误差值和实际值. 这个对比表明它们非常接近一致.

其间,拉普拉斯在 1810 年的一篇文章中给出了正态规律的一个新的理论推导. 拉普拉斯的结果建立在今天被称作中心极限定理的基础上,这是他对上世纪棣莫弗用二项式定理的计算的推广. 事实上,在假设由每次观测产生的误差是大致相等于 $-n, -n+1, -n+2, \cdots, -1, 0, 1, \cdots, n-2, n-1, n$ 的条件下,拉普拉斯证明了对较大的 s,由 s 次独立的观测产生的误差之和处于

$$-2T\sqrt{\frac{n(n+1)s}{6}} \quad 与 \quad 2T\sqrt{\frac{n(n+1)s}{6}}$$

之间的概率等于 $(2/\pi)\int_0^T e^{-x^2}dx$. 拉普拉斯也在误差概率的更一般的情形下导出了相似的结果. 不管怎样,拉普拉斯的工作在广泛的各式各样的情况下很快便把 $y = Ae^{-kx^2}$ 确立成了代表误差分布或更一般的概率分布的函数.

拉普拉斯把上述材料放进了 1812 年出版的他的主要著作《分析概率论》(Théorie analytique des probabilités) 中. 在这本书中拉普拉斯收集了到那时为止的所有的概率论发展的资料,开篇便定义一个事件的概率为“当没有什么能使我们相信一种情况会比任何其他情况更应该出现时,有利于它的情况的个数与可能情况的个数之比.”[64] 因此他也把中心极限定理的叙述和证明放在其中,还有这个定理对慧星轨道趋向问题的应用,这是他思考多年的问题. 另外,他还处理了概率论对一些专题的应用,如像对保险,人口统计,决策论和证据可信度. 事实上,拉普拉斯的观点正是,通过概率论,数学能够对社会科学发挥影响,就好像微积分是使物理学科数学化的主要工具一样.

16.5.3　统计学与社会科学

正态曲线 $y = Ae^{-kx^2}$ 首先是由棣莫弗从二项实验的概率计算中发展出来的,然后在对测量的极小化中显出了重要性. 在高斯和后来的拉普拉斯的公式化陈述中,它实际上代表了误差的分布. 但在 19 世纪中叶,这条曲线却被阿道夫 · 凯泰勒特(1796—1874)(图 16.11) 掌握成一把开启他的“平均人”概念的钥匙. 凯泰勒特是一位比利时数学家,天文学家,气象学家,社会学家,还是个统计学家. 在他汇编的统计资料中,不仅涵盖了像身高,体重这些物理特征还包括了“伦理特征”,如个体的犯罪倾向,或是酗酒倾向;由此他提出可以建立在给定时间,给定社会中的代表性人物这种概念,即平均人. 当然有按不同年龄,不同国家和不同阶级的不同平均男人(或者女



图 16.11　比利时邮票上的凯泰勒特.

人).这个概念的目的是将其用作理顺人们中各种各样的差异并在某种程度上揭示出社会的正常规律,即一种"社会物理学".

凯泰勒特注意到他所收集的许多特征可以借助于正态曲线来描绘.就是说存在一个均值以及与均值的"误差",它们以测量误差一样按相同的方式分布.1846年他写信给萨克森·科堡的大公爵表达了他在以这种方式使用正态曲线中的看法.他说,倘若一个人要去复制一千个某特定的塑像,这些复制品自然地有极多的误差,但是实际上这些误差却是以一种非常简单的方式组合成的.他写到,事实上"已经做过实验,是的,确实如此,上千个复制品已经被测量过 … 在全部情形中它们与原来的差别很小.这些复制品甚至像活的一样."[65] 凯泰勒特的"塑像"是些苏格兰士兵;他曾汇集了其中5732个的胸围,看出它们正态地分布在大约等于40英寸的均值周围.他的结论是,因为若性质瞄定的是一个理想类型的话,那么测量值就像它们可能的那样分布,一定是这样的.因此,他的分布表明了存在一个"平均的"苏格兰士兵而与这个平均的偏离只不过是由于各种偶然因素的组合造成的.

他在这种和其他的情况中使用正态曲线时,凯泰勒特的偏离平均值的单位是"可能误差"而不是现在通用的"标准偏离".一个数据点是离均值的一个可能误差是指,如果它有25或75的百分度.那就是,在一个正态分布中,一个特定值很可能在平均值的一个可能误差之内.这种对偏离的度量在世纪初已因与误差理论相关而被引用过.

确实有许多人不同意凯泰勒特的处处寻找正态分布的方案;但正态分布的思想在许多涉及统计的论证中成了中心.佛朗西斯·高尔顿(1822—1911)是一位英国统计学家,他运用凯泰勒特的思想试图以观察变异遗传的方式把达尔文的进化论数学化.他在1875年进行的一个实验是关于一种特定种类的甜豌豆的体积.他把七种不同大小的豌豆种子分成个数相等的组并研究它们的后代.原来每组的后代的大小是正态分布的,但在每组中的变化即数据的分布本质上是相同的.进一步他发现每组的平均值与父系种子的体积的不同.事实上平均值与父系种子的平均值线性相关,这条直线的斜率为三分之一.换句话说,第二代"退回"到总平均值.因而高尔顿开创了**回归**的统计研究(或像最初称呼的那样叫逆转).有趣的是,为计算这条回归线的斜率,高尔顿径直地画上这些点,然后再估算连接这些点的"最佳"直线的斜率.

高尔顿进行了涉及遗传的许多其他的研究,包括一个儿童身高与他们父母身高的关联的范围广泛的研究,他发现了类似的结果.于是他意识到在许多情形中,他所考虑的这两个变量是"相互关联的",就是说,当两个数据组都以可能误差单位来度量时,每个的回归线斜率与另一个是相同的.这个共同的斜率,或说大体上近似的斜率便是**相关系数**,可以用它来衡量两个变量间关联性的强度.当然,高尔顿知道一个强关联不一定意味着一个因果关系.

其他一些英国的统计学家力图澄清和推广高尔顿的工作并且认真地对待那些藏在代表各种分布的正态曲线的现象背后的基本哲学.弗朗西斯·埃奇沃斯(1845—1926)评论在用于天文学中的**观测**和在社会科学中收集而成的**统计**这两个概念之间的差别时说:"观测和统计作为在均值周围组成的数量时是一致的;它们的不同在于,其中观测的均值是真实的而统计的均值是虚构的.观测的均值是个原因,也可以说是个源头,从那里流出各种误差.统计的均值是一种描述,是整个群体的代表量,这个群体的最佳代表,这个量它 … 极小化了那些不可避免要加进这种实践中的误差."[66] 因此埃奇沃斯认识到在社会科学中使用统计不像在天文学中使用测量的误差理论那样,它不具有"客观"特性.尽管如此,他还是竭尽全力来发展这个新工具.

例如,埃奇沃斯在使用正态曲线来探索显著差时是非常保守的.事实上他用作正态曲线 $y=\frac{1}{c\sqrt{\pi}}e^{-(x^2/c^2)}$ 的偏差单位是模 c 而不是标准偏差 $c/\sqrt{2}$,他一般只得出这样的结论,即两个均值

的差是显著的是指它超过了两倍的模.这等价于使用了显著水平为 0.005 的一个双侧检定,比今天常用的 0.05 或 0.01 的水平来说是个非常强的要求.

在 19 世纪的最后十年间,两位英国的统计学家卡尔·皮尔逊(1857—1936)和他的学生乔治乌德尼·于勒(1871—1951)作了深入的工作来表明如何运用统计对许多个量之间的关系得出确定的结论.皮尔逊不仅在 1893 年引进了“标准偏差”同时也发展出了χ^2-统计法,它是一种度量两个量之间关系的方式.而于勒则指明如何利用在本质上是高斯最小二乘法的技巧来求出最佳直线的方式,以此计算回归方程.但是这些统计学家的所有方法都只是被设计来指明已经列出的量之间的关系.而今天统计学的一个显著应用是在实验程序的设计和分析之中,譬如说,它能使得农民用来确定对庄稼生产所用不同类型肥料的有效性,医生用来确定对某个特定疾病的不同处理方法的效力.这些方法首先是在 1930 年代通过罗纳尔德·费希尔(1890—1962)的工作发展起来的,他同时还创立了使用方便的显著性检验.有关运用统计过程的辩论却一直延续到今天.[67]

习　题

柯西的问题

1. 证明柯西的定理:如果$\lim\limits_{x\to\infty} f(x+1)-f(x)=\infty$,则$\lim\limits_{x\to\infty} f(x)/x=\infty$.
2. 利用习题 1 的定理证明$\lim\limits_{x\to\infty}\dfrac{a^x}{x}=\infty$及$\lim\limits_{x\to\infty}\dfrac{\log x}{x}=0$.
3. 证明$\dfrac{2}{\pi}\displaystyle\int_0^\infty \frac{x^2}{t^2+x^2}\mathrm{d}t=|x|$.
4. 利用连续性的现代定义和柯西的三角函数恒等式

$$\sin(x+\alpha)-\sin x=2\sin\frac{1}{2}\alpha\cos\left(x+\frac{1}{2}\alpha\right),$$

证明 $\sin x$ 对任一个 x 的值为连续.
5. 证明下面柯西的定理:如果对充分大的 x 值$f(x)$为正,并且当 x 无限增大时$f(x+1)/f(x)$收敛于 k,则$[f(x)]^{1/x}$当 x 无限增大时也收敛于 k.
6. 利用习题 5 的定理证明$\lim\limits_{x\to\infty}x^{1/x}=1$.
7. 用柯西准则证明级数$1+\dfrac{1}{1!}+\dfrac{1}{2!}+\dfrac{1}{3!}+\cdots$收敛.
8. 证明如果序列$\{a_i\}$收敛于 a,并且如果 f 连续,则序列$\{f(a_i)\}$收敛于$f(a)$.
9. 证明级数$\{u_k(x)\}$, $u_1(x)=x$, $u_k(x)=x^k-x^{k-1}$,其中 $k>1$,在 $x=1$ 的一个邻域中满足柯西的定理 6-1-1 的假定条件,但不满足此定理的结论.分析柯西对此种情形的证明以看出它的错误之处.
10. 用习题 4 中的三角公式证明正弦函数的导数是余弦函数.
11. 令 $a^i=1+\beta$,运用柯西的导数定义证明 $y=a^x$ 的导数是$y'=a^x/\log_a \mathrm{e}$.
12. 证明柯西在他的关于导数的定理中使用的代数结果:如果 $A<\dfrac{a_i}{b_i}<B$, $i=1,2,\cdots,n$,则

$$A<\frac{\sum_{i=1}^n a_i}{\sum_{i=1}^n b_i}<B.$$

13. 证明如果 f 在$[a,b]$上连续且如果 $a=x_0<x_1<\cdots<x_n=b$ 是$[a,b]$分成子区间的一个划分,则和

$$f(x_0)(x_1-x_0)+f(x_1)(x_2-x_1)+\cdots+f(x_{n-1})(x_n-x_{n-1})$$

等于$(b-a)f(x_0+\theta(b-a))$,其中 θ 在 0 与 1 之间.

14. 设 $f(x) = x^2 + 3x$,定义在$[1,3]$上.分$[1,3]$为8个子区间并确定那个满足习题13中性质的那个 θ 值.

波尔查诺的问题

15. 完成对最小上界准则的波尔查诺的证明:证明所构造出的序列的收敛值 U 是所有具有性质 M 的数的最小上界.
16. 设 A 为(3/5,2/3)之中那些小数展开式中只含有有限个0和有限个6而没有其他整数的数的集合.求出 A 的最小上界.
17. 设 M 为性质:$x^3 < 3$.由于这个性质不属于所有的 x 但却属于所有小于1的 x,故而这个性质满足波尔查诺最小上界定理的条件.从数 $V = 1 + 1$ 开始,这里比它小的数 x 不全满足 M,运用波尔查诺的证明方法构造出$\sqrt[3]{3}$的一个近似值,准确到小数点后三位.

傅里叶的问题

18. 证明 $\varphi(x) = \alpha e^{mx}, \psi(y) = \beta \cos ny$,它们是$\frac{\varphi(x)}{\varphi''(x)} = -\frac{\psi(y)}{\psi''(y)} = A$ 的解,其中 $m^2 = n^2 = 1/A$.最后得出 $v = ae^{-nx}\cos ny$ 是$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0$的解.
19. 证明$\int_0^\pi \sin mx \sin nx \mathrm{d}x = \begin{cases} 0 & \text{如果 } m \neq n \\ \pi/2 & \text{如果 } m = n. \end{cases}$
20. 计算一个函数 $\varphi(x)$ 的傅里叶余弦级数的系数 b_i.就是说,如果$\frac{1}{2}\pi\varphi(x) = b_0 + b_1\cos x + b_2\cos 2x + b_3\cos 3x + \cdots$,确定系数 b_i.
21. 利用傅里叶的积分法计算 $\varphi(x) = \frac{1}{2}x$ 的傅里叶级数,这个函数在正文中引用过.用其它已知的级数和检验此结果对 $x = \pi/2$ 的正确性.

有关连续性的问题

22. 考虑黎曼函数 $f(x) = \sum_{n=1}^{\infty}\frac{\varphi(nx)}{n^2}$,它定义在$[0,1]$上,其中 $\varphi(x)$ 等于 x 减去最靠近它的整数,但如果有两个整数等距地靠近它则 $\varphi(x)$ 为零.证明除去形如 $x = p/2n$,p 与 n 互素的无穷多个点外$f(x)$为连续.
23. 在加上级数一致收敛的假定下,证明关于连续函数项级数和的连续性的柯西定理.
24. 设 $u_k = \frac{1}{k(k+1)}$ 及 $v_k(h) = u_k + \frac{2h}{((k-1)h+1)(kh+1)}$,其中 h 为一个正的实变量.证明 $\lim_{h\to 0} v_k(h) = u_k$,证明$\sum u_k$ 收敛,并证明对充分小的 h,$\sum v_k(h)$ 收敛.另外再证明$\lim\limits_{h\to 0}\sum v_k(h) \neq \sum u_k$.

分析算术化的问题

25. 设 R 为如戴德金用分割定义的实数集.证明此集合具有基本的连续性的属性.即证明当 R 被分成两类 A_1, A_2,使得 A_1 中每个实数小于 A_2 中每个实数,则恰好存在一个实数 α,它或是 A_1 中最大的数或是 A_2 中最小的数.
26. 在由分割概念定义的戴德金的实数集 R 中定义一个自然的序 $<$.就是说,对给定的两个分割 $\alpha = (A_1, A_2)$ 与 $\beta = (B_1, B_2)$,定义 $\alpha < \beta$.证明这个序性在 R 上所满足的基本性质与它在有理数集上所满足的是相同的.
27. 定义戴德金分割的和.证明对任意两个分割 α 与 β 有 $\alpha + \beta = \beta + \alpha$.
28. 利用戴德金分割同时用康托尔的基本序列证明如下定理:每个实数的有界递增序列有极限.哪一种证明更容易些?
29. 证明如果$\{a_i\}$和$\{b_i\}$均为基本序列,并且$\{b_i\}$定义的不是极限0,则$\{a_i/b_i\}$也是基本序列.
30. 定义两个基本序列 $A = \{a_i\}$ 和 $B = \{b_i\}$ 的乘积 AB 为由积$\{a_ib_i\}$构成的序列.证明这个定义是有意义的,并且

如果 $AB = C$ 则 $B = C/A$,这里的除法如习题 29 中所定义的.

31. 显式确定一个点集 P,它的第一和第二导出集 P' 及 P'' 不同于 P 而且也互不相同.

32. 在 1890 年康托尔给出了下述定理的第二个证明,即区间(0,1) 中的实数不能与自然数集一一对应. 如果存在此区间中所有实数的一个排列 $\gamma_1,\gamma_2,\gamma_3,\cdots$. 把每个这样的数字写成无限小数形式:

$$\gamma_1 = 0.a_{11}\,a_{12}\,a_{13}\cdots,$$
$$\gamma_2 = 0.a_{21}\,a_{22}\,a_{23}\cdots,$$
$$\gamma_3 = 0.a_{31}\,a_{32}\,a_{33}\cdots.$$

现定义一个数 b,选取 $b = 0.b_1\,b_2\,b_3\cdots$,使得 $b_1 \neq a_{11}, b_2 \neq a_{22}, b_3 \neq a_{33},\cdots$. 证明 b 不可能在原来的排列中,从而不存在这样的排列.

复分析的问题

33. 用留数明晰地证明 $\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}\mathrm{d}x = \frac{\pi}{\mathrm{e}}$.

34. 用留数证明 $\int_0^{\infty} \frac{\mathrm{d}x}{1+x^6} = \frac{2\pi}{3}$.

35. 设复变函数 $w(z)$ 由和 $u(x,y) + \mathrm{i}v(x,y)$ 给出. 假定它满足了柯西 – 黎曼方程,即 $\partial u/\partial x = \partial v/\partial y, \partial v/\partial x = -\partial u/\partial y$. 证明导数 $\mathrm{d}w/\mathrm{d}z = \partial u/\partial x + \mathrm{i}\partial v/\partial x$.

向量分析的问题

36. 设曲面 S 由三个参数方程 $x = x(p,q), y = y(p,q), z = z(p,q)$ 定义. 几何地证明曲面元 $\mathrm{d}S$ 可以写成下面形式:

$$\mathrm{d}S = \left[\left(\frac{\partial(y,z)}{\partial(p,q)}\right)^2 + \left(\frac{\partial(z,x)}{\partial(p,q)}\right)^2 + \left(\frac{\partial(x,y)}{\partial(p,q)}\right)^2\right]^{1/2}\mathrm{d}p\mathrm{d}q.$$

37. 证明如果 $\sigma = A\mathrm{i} + B\mathrm{j} + C\mathrm{k}$ 为一向量场,满足 $\mathbf{div}\,\sigma = 0$,则 $\sigma = \mathbf{curl}\,\tau$,$\tau$ 为某个向量场.

38. 利用斯托克斯定理证明,如果 $\mathbf{curl}\,\sigma = 0$,则 $\int_C \sigma\cdot\mathrm{d}\boldsymbol{r}$ 与曲线 C 无关,只依赖于它的端点. 相似地,证明如果 $\mathbf{div}\,\sigma = 0$,则 $\int_S \sigma\cdot\mathrm{d}\boldsymbol{a}$ 与曲面 S 无关,只依赖于它的边缘曲线.

统计的问题

39. 对独立变量 x 给出了四个观测值为 $x_1 = 2.0, x_2 = 4.0, x_3 = 5.0$ 及 $x_4 = 6.0$ 及相应的对非独立变量 y 的四个观测值 $y_1 = 2.5, y_2 = 4.5, y_3 = 7.0$ 及 $y_4 = 8.5$;用最小二乘法确定表示此测量关系的最佳线性函数 $y = ax + b$ 中的常数 a 与 b.

40. 已知在正态曲线 $y = \frac{1}{c\sqrt{\pi}}\mathrm{e}^{-(x^2/c^2)}$ 中确定一个标准偏差的 x 值在此曲线的拐点上. 证明这个值由 $x = c/\sqrt{2}$ 给出.

41. 证明在具标准偏差 σ 的一个正态分布中,如果一个测量值离均值近似于 0.657σ,则其离均值为一个可能误差,如果它离均值为 $\sqrt{2}\sigma$,则测量值离均值为一个模(习题 40 中的 c 值).(应用正态曲线表解此问题.)

讨论题

42. 柯西说,一个无理数是各种趋向于它的分数的极限,他的意思是什么?柯西所理解的词"无理数",或者甚至词"数"是什么?

43. 请解释柯西对在一个区间上的连续性定义与通常在一个点上连续性的现代定义之间的差别. 是否一个满足柯

西定义的函数也满足在区间上每点都连续的现代定义?是否一个在区间中每点满足现代定义的函数也满足柯西的定义?

44. 详细设计一个讲授一致收敛性的教案,它以柯西的不正确定理及其证明作为开场.

45. 请注意,柯西没有像现今许多教科书所做的那样,使用罗尔定理来证明中值定理.讨论一下柯西的处理办法与标准的处理办法之间的优缺点.

46. 比较欧拉与黎曼对柯西 - 黎曼方程的推导.哪一个对复分析课程来说是个更好的引导?

文献和注解

关于 19 世纪分析学历史的最好的现代著作有:Ivor Grattan-Guinness 编纂的 *From the Calculus to Set Theory*(London:Duckworth,1980),它收集了专家们的各种专题的文章;Umberto Bottazzini 的 *The Higher Calculus:A History of Real and Complex Analysis from Euler to Weierstrass*(New York:Springer,1986);Ivor Grattan-Guinness 的 *The Development of the Foundations of Mathematical Analysis from Euler to Riemann*(Cambridge:MIT Press,1970).

1. Abel,"Investigation of the series $1+\frac{m}{1}x+\frac{m(m-1)}{1.2}x^2+\cdots$," *Crelle's Journal* 1(1826),311 - 339,p.316.
此文的一部分,包括这个引述,被翻译在 G.Birkhoff,*A Source Book in Classical Analysis* (Cambridge:Harvard University Press,1973),p.68 - 70.此书包含了 19 世纪期间许多最重要的分析方面的文章.

2. Sylvestre Lacroix,*An Elementary Treatise on the Differential and Integral Calculus*,translated by Charles Babbage,George Peacock,and John Herschel (Cambridge:J.Deighton,1816),p.2.

3. 同上,p.5.

4. Oystein Ore,*Niels Henrik Abel:Mathematician Extraordinary*(New York:Chelsea,1974),p.147.

5. Lacroix,preface.

6. Cauchy,*Cours d'analyse de l'école royale polytechnique*,重印于 Cauchy,*Oeuvres complète d'Augustin Cauchy*(Paris:Gauthier-Villars,1882 -)(2),3,p.19.(所有进一步对 *Cours d'analyse* 其他内容的参考页码都是根据此全集,关于 Cauchy 的所有其它论述也都将参照了此全集 *Oeuvres*.)对柯西在微积分方面的最好的材料是 Judith V.Grabiner,*The Origins of Cauchy's Rigorous Calculus*(Cambridge:MIT Press,1981).本章第一节的许多细节都取自此书.对柯西和其他一些人在连续性观念方面工作的一个简短概述可参考 Judith V.Grabiner,"Who gave you the epsilon?Cauchy and the origins of rigorous calculus,"*American Mathematical Monthly* 90(1983),185 - 194.而对有关导数概念发展方面的评述可见于 Judith V.Grabiner, "The changing concept of change:The derivative from Fermat to Weierstrass,"*Mathematics Magazine* 56(1983),195 - 203.

7. Cauchy,*Cours d'analyse*,p.54.

8. Lagrange,*Théorie des fonctions analytique*,收录于 *Oeuvres de Lagrange*(Paris:Gauthier-Villars,1867—1892),vol.9,p.28. Quoted in *Grabiner,Cauchy's Rigorous Calculus*,p.95.

9. Cauchy,*Cours d'analyse*,p.43.

10. S.B.Russ,"A Translation of Bolzano's Paper on the Intermediate Value Theorem,"*Historia Mathematica* 7(1980),156 - 185,p.159.对于波尔查诺工作的更多的讨论可见于 I.Grattan-Guinness,"Bolzano,Cauchy and the'New Analysis'of the Early Nineteenth Century,"*Archive for History of Exact Sciences* 6(1970),372 - 400.Grattan-Guinness 宣称柯西从波尔查诺那里获取了连续性和收敛性定义的中心思想,参见 H.Freudenthal,"Did Cauchy Plagiarize Bolzano?"*Archive for History of Exact Sciences* 7(1971),375 - 392,它的观点是不同的.

11. 同上,p.162.

12. Cauchy,*Cours d'analyse*,p.114.

13. 同上,p.116.

14. 引自 A.J.Franco de Oliveira,"Anastácio da Cunha and the Concept of Convergent Series,"*Archive for History of Exact*

Sciences 39(1988),1 - 12,p.4.对达·库尼亚更多的材料见 João Filipe Queiró,"José Anastácio da Cunha: A Forgotten Forerunner,"*The Mathematical Intelligencer* 10(1988),38 - 43 以及 A.P.Youschkevitch, "J.A.da Cunha et les fondements de l'analyse infinitésimale,"*Revue d' Histoire des Sciences* 26(1973),3 - 22.

15. Russ,"Translation of Bolzano's Paper,"p.171.
16. 对柯西的错误证明的论述取自 V.Frederick Rickey 未发表的文稿 *Using History in Teaching Calculus*.
17. Cauchy, *Resumé des leçons données a l' école polytechnique sur le calcul infinitésimal*,收录于 *Oeuvres* (2),4,p.44.
18. Lacroix, *Elementary Treatise*, p.179.
19. Cauchy, "Mémoire sur l'intégration des équations lineares aux différentielles partielles et a coefficients constantes,"in *Oeuvres*,(2),1,275 - 357,p.354.
20. 同上,p.334.
21. Cauchy, *Resumé*, pp.122 - 123.
22. 同上,p.125.
23. 引自 Grattan-Guinness, *From the Calculus to Set Theory*, p.158.
24. 同上,p.153.
25. 引自 Grattan-Guinness, *Foundations of Mathematical Analysis*, p.104.
26. 引自 Bottazzini, *Higher Calculus*, p.244.
27. 引自 Ann Hibner Koblitz, *Sofia Kovalevskaia: Scientist, Writer, Revolutionary*(Boston: Birhkäuser, 1983), p.197. 本书提供了柯瓦列芙斯卡娅的详细的非专业性的传记,同时也对她的数学有所描述.她的数学工作的更详细的处理见 Roger Cooke, *The Mathematics of Sonya Kovalevskaya*(New York: Springer-Verlag, 1984).
28. E.Heine,"Die Elemente der Functionenlehre,"*Journal für die Reine und Angewandte Mathematik* 74(1872),172 - 188, p.184.
29. Dedekind, *Continuity and Irrational Numbers*, Wooster Beman 译,收录于 Dedekind, *Essays on the Theory of Numbers* (La Salle, Ill: Open Court, 1948), pp.1 - 2.
30. 同上,p.9.
31. 同上,p.11.
32. 同上,pp.12 - 13.
33. 同上,p.15.
34. 同上,p.22.
35. 引自 Bottazzini, *Higher Calculus*, p.277.
36. 同上,p.278.
37. 引自 Joseph Dauben, *Ceorg Cantor: His Mathematics and Philosophy of the Infinite*(Princeton: Princeton University Press, 1979), p.49. Dauben 写的传记给出了对康托尔工作的详细研究以及它是如何与他那个时代的数学与哲学相关联的.
38. Cantor, *Contributions to the Founding of the Theory of Transfinite Numbers*, translated by Philip Jourdain (Chicago: Open Court, 1915), p.85.
39. 同上,p.86.
40. 引自 Gregory H.Moore, "Towards a History of Cantor's Continuum Problem,"in David Rowe and John McCleary, eds., *The History of Modern Mathematics* (San Diego: Academic Press, 1989), vol.1, 79 - 121, p.82. Rowe 与 MeCleary 的两卷本著作刊出了在 1988 年举行的关于 19 世纪数学的会议的记录.这些文章很值得一读.
41. Dedekind, *The Nature and Meaning of Numbers*, translated by Wooster Beman, in Dedekind, *Theory of Numbers*, p.45.
42. 同上,p.50.
43. 同上,p.53.
44. 同上,p.63.
45. 同上,p.64.

46. Wessel, "On the Analytical Representation of Direction," translated in David Smith, *A Source Book in Mathematics* (New York: Dover, 1959), 55 – 66, p.58.

47. 引自 Bottazzini, *Higher Calculus*, p.156, from the letter published in Gauss, *Werke* (Göttingen, 1866), vol.8, pp.90 – 92.

48. Cauchy, *Cours d'analyse*, p.154.

49. Cauchy, *Mémoire sur les intégrales définies prises entre des limites imaginaires* (Paris, 1825), pp.42 – 43, Birkhoff译, *Source Book*, p.33.

50. 同上, p.44; p.34.

51. Riemann, *Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse*, parts of which are translated in Birkhoff, *Source Book*, 48 – 50, p.48.

52. 同上, p.49.

53. Riemann, "Theorie der Abel'sche Funktionen," *Crelle's Journal* 54(1857), parts of which are translated in Birkhoff, *Source Book*, 50 – 55. p.51.

54. 同上, pp.52 – 53.

55. 同上,

56. 关于散度定理以及格林定理和斯托克斯定理可以在下文中找到: Victor J. Katz, "The History of Stokes' Theorem," Mathematics Magazine 52(1979), 146 – 156.

57. Stokes, *Mathematical and Physical Papers* (Cambridge: Cambridge University Press, 1905). Vol.5, p.320.

58. Legendre, "Sur la Méthode des moindres quarrés," in Legendre, *Nouvelles Méthodes pour la détermination des orbites des cometes* (Paris, 1805), translated in Smith, Source Book, 576 – 579, p.576.

59. 同上, p.577.

60. 更详细的见 R.L. Plackett, "The discovery of the method of least squares," *Biometrika* 59(1972), 239 – 251. 此文重印在 M.G. Kendall and R.L. Plackett, eds., *Studies in the History of Statistics and Probability*, vol.2 (New York: Macmillan, 1977), 279 – 291. 此卷及其同集著作 E.S. Pearson and M.G. Kendall, eds., *Studies in the History of Statistics and Probability*, vol.1 (Darien, Conn.: Hafner, 1970), 提供了关于概率论和统计学历史方面的文章汇编, 其中大部分文章最初出现于 *Biometrika*.

61. 更详细的讨论见 Steven C. Althoen and Renate McLaughlin, "Gauss-Jordan Reduction: A Brief History," *The American Mathematical Monthly* 94(1987), 130 – 142 和 Victor J. Katz, "Who is the Jordan of Gauss-Jordan?" *Mathematics Magazine* 61(1988), 99 – 100.

62. 引自 Stephen M. Stigler, *The History of Statistics* (Cambridge: Harvard University Press, 1986), p.94. 这个著作提供了统计学从起源经 1900 年的发展的详细研究, 它是正文 16.5.3 中讨论材料的来源.

63. Gauss, *Theoria motus corporum celestium* (Hamburg: Pertheset Besser, 1809), translated by C.H. Davis as *Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections* (Boston: Little, Brown, 1857), p.258. 对高斯及最小二乘法的进一步信息见 O.B. Sheynin, "C.F. Gauss and the Theory of Errors," *Archive for History of Exact Sciences* 20(1979), 21 – 72 和 William C. Waterhouse, "Gauss's First Argument for Least Squares," *Archive for History of Exact Sciences* 41(1991), 41 – 52. 对一位也曾对最小二乘法做出贡献的早期美国数学家 Robert Adrain (1775—1843) 的工作的讨论见 Jacques Dutka, "Robert Adrain and the Method of Least Squares," *Archive for History of Exact Sciences* 41 (1991), 171 – 184.

64. Laplace, *Théorie analytique des probabilités* (Paris: Courcier, 1812), p.181.

65. Adolphe Quetelet, *Letters Addressed to H. R. H. the Grand Duke of Saxe Coburg and Gotha, on the Theory of Probabilities, as Applied to the Moral and Political Sciences*, translated by O.G. Downes (London: Charles and Edwin Layton, 1849), p. 136.

66. Francis Ysidro Edgeworth, "Observations and statistics: An essay on the theory of errors of observation and the first principles of statistics," *Transactions of the Cambridge Philosophical Society* 14(1885), 138 – 169, p.139. 对这篇文章在 Stigler, *History of Statistics*, Chapter 9 中有所讨论.

67. 关于 19 世纪统计学的更多信息参见 Stigler, *History of Statistics* 和 Theodore M. Porter, *The Rise of Statistical Thinking*(Princeton:Princeton University Press,1986). 关于 20 世纪的发展的信息见 Gerd Gigerenzer,et al., *The Empire of Chance: How probability changed science and everyday life* (Cambridge:Cambridge University Press,1989) 和 Lorenz Krüger,Lorraine Daston,and Michael Heidelberger,eds., *The Probabilistic Revolution*(Cambridge:MIT Press, 1987).

19 世纪分析学概览

1736—1813	约瑟夫·路易斯·拉格朗日(Joseph Louis Lagrange)	曲面积分
1744—1787	约瑟·阿拉斯塔西奥·库尼亚(José Anastácio da Cunha)	收敛性的定义
1745—1818	卡斯帕·韦塞尔(Caspar Wessel)	复数的几何表示
1749—1827	皮尔斯·西蒙·拉普拉斯(Pierre Simon Laplace)	分析概率论
1752—1833	阿德瑞恩·玛瑞·勒让德(Adrien Marie Legendre)	最小二乘法
1765—1843	弗朗索瓦·西尔维斯特·拉克鲁瓦(Sylvestre-François Lacroix)	微积分教科书
1768—1830	约瑟夫·傅里叶(Joseph Fourier)	傅里叶级数
1777—1855	卡尔·费里德里希·高斯(Carl Friedrich Gauss)	复变函数、正态分布
1781—1848	伯恩哈德·波尔查诺(Bernhard Bolzano)	连续性与收敛性
1789—1857	奥古斯汀·路易斯·柯西(Augustin-Louis Cauchy)	分析的严谨化、复分析
1796—1874	阿道夫·凯泰勒特(Adolphe Quetelet)	正态曲线
1801—1861	米哈伊尔·奥斯特罗格拉茨基(Mikhail Ostrogradsky)	散度定理
1802—1829	尼尔斯·亨利克·阿贝尔(Niels Henrik Abel)	收敛性的新思想
1805—1859	彼得·勒热纳·狄利克雷(Peter Lejeune-Dirichlet)	傅里叶级数的收敛性
1815—1897	卡尔·魏尔斯特拉斯(Karl Weierstrass)	一致收敛性
1819—1903	乔治·斯托克斯(George Stokes)	向量分析
1821—1881	艾杜尔德·海涅(Eduard Heine)	连续性的新思想
1822—1911	弗朗西斯·高尔顿(Francis Galton)	回归与相关
1826—1866	乔治·伯恩哈德·黎曼(Georg Bernhard Riemann)	积分、拓扑的开端
1831—1879	理查德·戴德金(James Clerk Maxwell)	戴德金分割、自然数的公理
1831—1916	詹姆士·克拉克·麦克斯韦(Richard Dedekind)	线积分
1845—1918	乔治·康托尔(Georg Cantor)	集合论
1845—1926	弗朗西斯·埃奇沃斯(Francis Edgeworth)	统计方法
1850—1891	索菲亚·柯瓦列芙斯卡娅(Sofia Kovalevskaya)	偏微分方程
1857—1936	卡尔·皮尔逊(Karl Pearson)	统计方法
1871—1951	乔治·乌迪尼·于勒(George Udny Yule)	统计方法

第17章 19世纪的几何学

“我愈来愈深信我们不能证明我们的几何具有必然性——至少不能用推理来证明.或许在另一个世界中,我们可能洞察空间的性质,而现在这是不能达到的.同时我们不能把几何与算术相提并论,因为算术是纯粹先验的,而几何却可以和力学相提并论.”

——摘自高斯1817年写给奥尔伯斯(Heinrich Olbers)(1758—1840)的一封信[1]

欲成为哥廷根大学的数学讲师,必不可少的一步即是给校教授团发表就职演说,这次演说旨在展示候选者将数学研究应用于更加普遍的学术问题的才华.黎曼为这次任职资格考核提交了三个备选题目,前两个与复变函数及三角级数的一些已完善的研究紧密相联.但是,高斯代表教授团选择了黎曼的第三个题目《关于几何基础的假设》.1854年6月10日,黎曼进行了演讲陈述.这次演说很少涉及到数学上的细枝末节,但却蕴涵了大量的关于几何学应如何发展的真知灼见,而这些问题,数学家们为之思索了一个多世纪.

18世纪后期,分析学得以居于主导地位,很大程度上归功于欧拉的工作,而柯西的努力使之持续到了19世纪初期.在此期间,其他很多人却忽视了纯几何学的重要性.将分析应用于几何学,却产生了各种新的、重要的几何思想.

尽管高斯在他学术生涯的早期即考虑到今天所谓的微分几何的方方面面,但直到他作为哥廷根天文台台长,对汉诺威王国具体进行勘测期间,他才最终阐明了曲面理论的观点.1827年,高斯发表了一篇精短论文来陈述这些观点,他提出将欧拉的理论引入曲面,并应用微积分的技巧来阐释曲面理论的一些基本概念,包括对曲面来说是内蕴的,并不依赖所处的三维空间的曲率的概念.

在有关曲面的著作中,高斯确立了曲率与曲面三角形角度之和的关系.这个关系被认为是与古老的平行公设问题紧密联系.而平行公设蕴涵平面三角形内角之和等于两直角的论断.在其晚年,高斯提到,他很早就相信平行公设是不能被证明的,如果用其它公设代替,将建立起新的、有趣的几何学体系.对物质世界来说,此体系的正确性仅能建立在实验的基础上.尽管如此,高斯从没发表有

关这个主题的观点.所以,直到 19 世纪 20 年代,罗巴切夫斯基(Nikolai Lobachevsky) 和 J. 波尔约(János Bolyai) 才各自独立的首次完整地发表了非欧几何的论文.

但是,在将近 40 年后,非欧几何的思想才得以引起数学界的注意.直到 1854 年黎曼及 1868 年 H.赫姆霍尔兹(Hermann von Helmholtz) 的有关任意维几何流形的一般观点的著作的发表,这些有关几何学研究的新思想才站稳脚跟.不久以后,非欧几何的各种模型被引入到欧氏空间,这使数学界相信,从逻辑观点而言,非欧几何和欧几里得几何同样正确.而且,在我们所生活的世界里,欧氏几何的正确性,不再有显然的答案.

射影几何学也同样得到了发展,其奠基于帕斯卡(Pascal) 和德扎格(Desargues) 的早期的工作,而经过彭赛列(Jean-Victor Poncelet),沙勒(Michel Chasles) 和普吕克(Julius Plücker) 等一批数学家的努力得到了完善. 1871 年克莱因(Felix Klein) 使用度量的思想,将射影几何和非欧几何联系起来.第二年,他利用变换给出了几何学的一个新定义,此定义证明了射影几何和欧氏几何的关系,以及几何和那时正在成型的群论的联系.

19 世纪中期,几何学研究的对象仍未超过三维.但随着越来越广泛的使用分析和代数的方法,人们开始明确:不应该由于特殊原因,将几何的维数限制在恰能被实际感知的数目上.这样,好些数学家便将他们的公式和理论推广到 n 维,这里 n 可以取任何正整数.格拉斯曼(Hermann Grassmann) 于 1844 年率先尝试从几何观点深入研究 n 维向量空间.但不幸的是,格拉斯曼的工作,就像 J. 波尔约和罗巴契夫斯基那样,直到世纪末才受到重视.那时,佩亚诺(Giuseppe Peano) 给出有限维向量空间的一套公理体系,为高维几何的研究提供了基础,E.嘉当(Elie Cartan) 将格拉斯曼的研究应用到了微分形式的研究中.

随着许多新几何学科的产生,19 世纪晚期,许多数学家感到,就像分析学正在做的那样,是重新筑建整个学科的根基的时候了.欧氏几何的论证已经发现了缺陷,这给非欧几何的发展带来了一定问题.于是,希尔伯特(David Hilbert) 为欧氏几何提出了一套新的公理体系,这将有助于消除各种缺陷,并切实指明了为使新旧几何学均得以发展而必做的那部分假设.

17.1　微分几何学

1820 年至 1825 年,高斯领导了对汉诺威的勘测工作,引入了多种新方法来建立大地测量学,使之成为普遍接受的科学. 1827 年,高斯终于将他关于曲面的若干想法诉之以文,这个问题,他已思索了二十五年多.高斯在他的论文《空间曲面的一般研究》的摘要中写道:

> 尽管几何学者们对空间曲面的一般研究给予很大关注,其成果涵盖高等几何学的重要领域,这个主题仍然远未阐述详尽.可以这么说,迄今为止,这块富沃的土地仅被开垦了一小部分.通过寻求问题的解,来得到在微元保持不变的另一个曲面上,所给原曲面的所有表示.对于此项研究,若干年以前作者即倾注精力,探索出一种新途径.现在讨论的目的在于进一步启发新观点,并建立一些目前可以接受的新理论.[2]

1822 年,高斯解决了他自己提议哥本哈根科学会提出的一个挑战性的问题,即如何设立条件建立一个曲面共形地映到另一个曲面(使微元保持不变) 的映射.答案是对两个曲面进行参数表示时,函数需表示为复解析的(但那时高斯未使用复变函数理论).在寻求这个答案的过程中,高斯意识到曲面研究的中心问题是曲率问题,尤其是他解决了一直考虑的由曲面的解析描述计算出曲率的问题.

17.1.1 曲率的定义

在《一般研究》的开头，高斯给出了曲面曲率的概念. 他决定只涉及具有连续曲率的曲面或部分曲面，这类曲面(或部分)在所有点上都具有切平面. 这样，锥体的顶点将不再被考虑，因为在那点上不存在切平面和曲率. 因为球面是高斯的曲面"模型"，具有类似于平面圆周的常曲率，高斯决定由曲面上一点周围的区域与单位球面上对应点周围区域比较来定义出此点曲率. 对曲面 S 来说，曲率是一个局部性质，但显而易见的是，它随点而变. 在单位球面上，每点曲率均置为 1. 为了比较，高斯建立了一个从 S 到单位球面的映射 n(即今天所谓的高斯法映射)，使从球中心指向 $q = n(p)$ 的向量与 S 在 p 点的法向量(即 S 在 p 点的切平面的法向量)平行(图 17.1). 这个映射把曲面 S 上的有界区域 A 映到球面上的有界区域 $n(A)$. 于是高斯定义出 A 的**全曲率**(total curvature)为 $n(A)$ 的面积，同时在一点处他定义了更为重要的**曲率测度**(measure of curvature)概念为"用曲面上一元素面积去除其[全]曲率的比值，因此它代表了曲面和球面对应的无限小面积之比"[3]. 用现在的术语，高斯定义点 p 的曲率为：

$$k(p) = \lim_{A\to p}\frac{n(A)\text{ 的面积}}{A\text{ 的面积}},$$

这里极限是含有 p 的区域 A 收缩为 p 本身时取的. 从而可得区域 A 的全曲率是 $\int k\mathrm{d}\sigma$，$\mathrm{d}\sigma$ 是曲面面积元素，积分区域为 A.

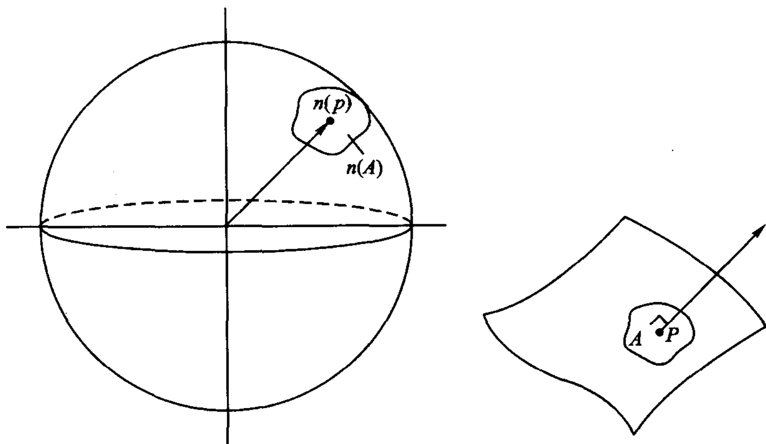


图 17.1 高斯的法映射.

现在的读者可能会注意到高斯曲率测度的定义至少存在两个问题. 第一，在任意一个曲面上，面积如何定义；第二，假设这一点已解决，我们如何知道，若极限存在，它是否和区域的收缩方式无关？高斯没有谈到这些问题. 事实上，他并没有将曲率定义为极限，而仅将之视为无穷小的比率. 他由几何直觉确信所下定义有意义. 例如，如果曲面 S 是半径为 r 的球面，则 $n(A)$ 的面积是 A(对任意区域 A)面积的 $1/r^2$ 倍. 因此，每点曲率均为 $1/r^2$. 类似地，如果 S 是一个平面，则对任意区域 A 来说，$n(A)$ 均为一个点. 它的面积为 0，曲率也为 0. 更让人吃惊的是，当 S 是圆柱面时，高斯同样确信

他的定义是正确的.在这种情况下,区域 A 的象集 $n(A)$ 只是球面上一条曲线,因而面积是 0,从而柱体曲率和平面曲率一样也是 0.后面将简要论述结论为何合理.

17.1.2 曲率和"绝妙的定理"

依据所给的曲面方程,高斯可以用他的定义计算出曲率.因为 S 在 p 点的切平面平行于单位圆在 $n(p)$ 的切平面,$n(A)$ 与 A 面积之比和 $n(A)$ 与 A 在 xy- 平面射影面积之比相等.因而,就一个其射影顶点为 (x,y),$(x+\mathrm{d}x, y+\mathrm{d}y)$,$(x+\delta x, y+\delta y)$ 的三角区域 A,高斯指出此三角形面积为 $\frac{1}{2}(\mathrm{d}x\delta y-\mathrm{d}y\delta x)$.类似地,如果函数 $X(x,y)$,$Y(x,y)$ 表示法向函数 n 与射影的合成,那么 $n(A)$ 相应的面积为 $\frac{1}{2}(\mathrm{d}X\delta Y-\mathrm{d}Y\delta X)$,则有

$$k=\frac{\mathrm{d}X\delta Y-\mathrm{d}Y\delta X}{\mathrm{d}x\delta y-\mathrm{d}y\delta x}.$$

现在留待高斯确定的是,如果曲面由方程 $z=z(x,y)$,$w(x,y,z)=0$ 或参数方程 $x=x(p,q)$,$y=y(p,q)$,$z=z(p,q)$ 定义,此分式的值是什么?对第一种表示法,由

$$\mathrm{d}X=\frac{\partial X}{\partial x}\mathrm{d}x+\frac{\partial X}{\partial y}\mathrm{d}y,\quad \mathrm{d}Y=\frac{\partial Y}{\partial x}\mathrm{d}x+\frac{\partial Y}{\partial y}\mathrm{d}y,$$

$$\delta X=\frac{\partial X}{\partial x}\delta x+\frac{\partial X}{\partial y}\delta y,\quad \delta Y=\frac{\partial Y}{\partial x}\delta x+\frac{\partial Y}{\partial y}\delta y,$$

高斯得出

$$k=\frac{\partial X}{\partial x}\frac{\partial Y}{\partial y}-\frac{\partial X}{\partial y}\frac{\partial Y}{\partial x}.$$

此式虽直接但杂乱,根据 z 的偏导数计算后重新表述为:

$$k=\frac{z_{xx}z_{yy}-z_{xy}^2}{(1+z_x{}^2+z_y{}^2)^2}.$$

高斯计算出当曲面由三变元方程表示和由参数方程表示时,k 的表述类似.从而,他得出了一系列优美的定理.

首先,他指出,根据两个经过 p 点的特殊截面曲线的曲率可以表示此点曲率测度.经过选择合适的坐标轴,在 $p=(0,0,0)$,$z_x(0,0)=z_y(0,0)=z_{xy}(0,0)=0$ 形式下重写了方程 $z=z(x,y)$.则 $z_{xx}(0,0)$ 和 $z_{yy}(0,0)$ 分别为经过 p 点的形成的所有曲线的极大曲率和极小曲率.这样,p 点曲率测度 $k(p)$ 为 $z_{xx}z_{yy}$,即正截口的两个极曲率的积.

其次,高斯证明了他的"绝妙的定理":曲率测度在曲面上是等距不变的,也即,如果对曲面进行保持距离的变换,曲率测度不会改变.为证明此定理,他采用曲面的参数表示形式,$x=x(u,v)$,$y=y(u,v)$,$z=z(u,v)$,设 $E=x_u^2+y_u^2+z_u^2$,$F=x_ux_v+y_uy_v+z_uz_v$ 和 $G=x_v^2+y_v^2+z_v^2$,并由 E,F,G 和它们对 u,v 的一阶,二阶偏导数,独立地得到曲率的表达公式.不难得到距离元 $\mathrm{d}s$ 自身也可被表示由这些量表示 $\mathrm{d}s^2=\mathrm{d}x^2+\mathrm{d}y^2+\mathrm{d}z^2=E\mathrm{d}u^2+2F\mathrm{d}u\mathrm{d}v+G\mathrm{d}v^2$.这样,曲率即可由距离元决定.于是,高斯提出了他的"著名"定理,即,如果一个曲面在另一个曲面上"展开",也即若存在从一个曲面到另一个的一一对应的函数来保持长度元不变,则两曲面对应点上的曲率测度总相等.例如,因为平面可在圆柱上展开,圆柱曲率和平面曲率相等,即为 0.高斯强调他的结论只是通往研究曲面的重要新途径的开端:

> 不止把曲面看作立体的边界,而是看作富有弹性但不可延伸的,消失了一个维数的立

体.[在这种方式下]曲面的性质一部分依赖我们已将其约化成的形式,另一部分则是绝对的,无论曲面被弯曲成何种形式,总保持不变.对后者的研究为几何学开辟了一块新奇富饶的领域,得出曲率测度和整曲率……从这个观点来看,平面和可在平面上展开的曲面,例如,圆柱曲面,圆锥曲面等等,均被认为本质上是相同的.[4]

最后,高斯证明了由曲面上的测地弧(长度最短的弧)形成的三角形的全曲率和三角形内角之和的重要关系.事实上,他计算出测地三角形的全曲率$\int k\mathrm{d}\sigma$等于$A+B+C-\pi$,这里,A,B,C是三角形三个角的量度.例如,在一个有正的常曲率的曲面上,每个测地三角形角度之和大于π,而在一个有负的常曲率的曲面上,每个测地三角形角度之和小于π.高斯的结论是对一个众所周知的结论的推广,这个结论是,在一个单位球面上,一个区域的全曲率等于其面积,这时一个由大圆弧(测地线)组成的三角形的三个角之和减去π等于此三角形的面积.

高斯关于曲面微分几何的论述,不仅自身意义重大,而且对未来的工作有着深远的影响.尤其是,他证明出三角形内角之和与曲面上内蕴几何的关系,有助于引向解决有关平行公设有效性的问题.而且,高斯用长度元表示曲面特征,而长度元又可由量E,F,G表示,被认为开n维流形一般理论的先河.大约30年后,黎曼发展了这个理论的许多重要内容.

17.2 非欧几里得几何

在18世纪,沙克里和朗贝特都曾尝试经由先假定公设错误,由此导出矛盾的途径来证明欧几里得平行公设.沙克里认为在探索这个问题上,他已取得成功,但朗贝特却认识到他的尝试失败了.在此问题上,两人均以综合法入手,试图使用欧几里得方法论来说明欧几里得假定了一条多余的公设.但到了19世纪,分析越来越多的被用来解决各式各样的问题,同样也为此问题提供了一条新思路.而且,颇为有趣的是,正是朗贝特的双曲函数为建立分析与一种新几何间的联系作出贡献,而这个联系朗贝特本人并未注意到.

17.2.1 陶里纳斯和对数球面几何

朗贝特曾指出锐角假设似乎可在虚半径球面上成立,但事实上是1826年,由陶里纳斯(Franz Taurinus,1794—1874)在一篇论文中明确建立了这种联系,陶里纳斯是一个善于独立思考的数学爱好者,他由球面三角学的一个公式入手.这个公式建立了半径为K的球面上,任意一球面三角形一内角和所有边的联系:

$$\cos\frac{a}{K}=\cos\frac{b}{K}\cos\frac{c}{K}+\sin\frac{b}{K}\sin\frac{c}{K}\cos A,$$

这里,三角形三边为a,b,c,对顶角分别为A,B,C(图17.2).用$\mathrm{i}K$替换K,也即使球面半径虚数化(无论以何种方式),陶里纳斯利用$\cos \mathrm{i}x=\cosh x$和$\sin \mathrm{i}x=\mathrm{i}\sinh x$,得到一个新公式:

$$\cosh\frac{a}{K}=\cosh\frac{b}{K}\cosh\frac{c}{K}-\sinh\frac{b}{K}\sinh\frac{c}{K}\cos A. \qquad (17.1)$$

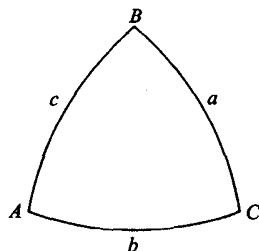


图17.2 半径为K的球面上的球面三角形.

陶里纳斯将由此公式定义的几何称为"对数球面几何",但他意识到这个几何不适用于平面.但是,对公式结果的探求,可以得出这个几何的一些特性.例如,若三角形为等边的($a=b=c$),公式变为:

$$\cosh\frac{a}{K} = \cosh^2\frac{a}{K} - \sinh^2\frac{a}{K}\cos A$$

或

$$\cos A = \frac{\cosh^2\frac{a}{K} - \cosh\frac{a}{K}}{\sinh^2\frac{a}{K}} = \frac{\cosh\frac{a}{K}\left(\cosh\frac{a}{K} - 1\right)}{\cosh^2\frac{a}{K} - 1} = \frac{\cosh\frac{a}{K}}{\cosh\frac{a}{K} + 1}.$$

因为 $\cosh\frac{a}{K} > 1$,从而可得 $\cos A > 1/2$,因而 $A < 60°$.换言之,在此几何中,正三角形内角之和小于 180°.另一方面,容易看出,当任意一边变小或半径 K 变大,角 A 均趋近 60°,而这个几何也接近欧几里得几何.事实上,我们可以指出(通过幂级数展开),当取极限使 K 趋于 ∞,陶里纳斯的公式(17.1)简化为欧几里得余弦定理 $a^2 = b^2 + c^2 - 2bc\cos A$.

球面三角学的又一个重要的公式是:

$$\cos A = -\cos B\cos C + \sin B\sin C\cos\frac{a}{K}.$$

此公式建立了球面三角形边和角的关系.用 $\mathrm{i}K$ 替换 K,则变为对数球面几何上的公式:

$$\cos A = -\cos B\cos C + \sin B\sin C\cosh\frac{a}{K}. \tag{17.2}$$

当 $A = 0°, C = 90°$ 的特殊情况下,公式(17.2)简化为:

$$\cosh\frac{a}{K} = \frac{1}{\sin B}.$$

当然,在欧氏几何,角 C 为直角,角 A 为 0° 的三角形不存在.但是,沙克里由锐角假设得出了渐近线的概念.这样,陶里纳斯的三角形必须有两边是渐近的(图 17.3).于是,角 B 和第三边长度 a 由公式 $\sin B = \mathrm{sech}(a/K)$ 联系起来.形式上,此公式可改写为:

$$\tan\frac{B}{2} = \mathrm{e}^{-a/K},$$

这个公式后来成为罗巴切夫斯基理论的基础.

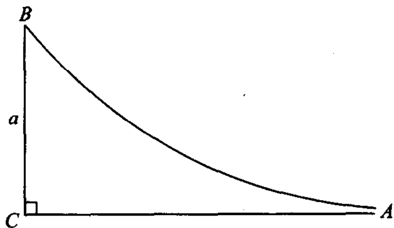


图 17.3　$A = 0°, C = 90°$ 的三角形.

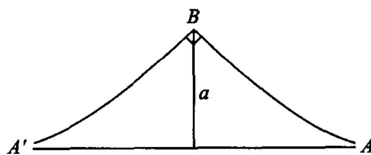


图 17.4　两个边角为 0 度的等边直角三角形.

公式(17.2)也表明,通过引顶垂线 a 将原等腰直角三角形分为两个直角三角形,则 a 和原三角形底角 A 的关系为 $\cosh\frac{a}{K} = \sqrt{2}\cos A$(图 17.4).从而当 $A = 0°$ 时,即三角形的两个侧边都是渐近地趋向斜边时,等腰直角三角形的顶垂线取极大值 h.在这种情况下 $\cosh\frac{h}{k} = \sqrt{2}$ 或 $K = \frac{h}{\ln(1+\sqrt{2})}$.

陶里纳斯进一步注意到三角形面积和它的亏量成比例(这一点朗贝特早就发现了),半径为 r

的圆其周长为 $2\pi K\sinh\frac{r}{K}$，面积为 $2\pi K^2(\cosh\frac{r}{k}-1)$. 值得注意的是，这些后来的结果和不久后在罗巴切夫斯基和波尔约的论文中涉及到的一些结论，在稍微早些的高斯的未公开的论文中已经存在. 但可能是由于高斯以他的严格的标准来看，认为自己未能证明出所有结论，所以从未直接发表有关这个主题的任何结论. 另一方面，高斯的联系曲面曲率和三角形亏量或超量的著作多多少少必然受到他的新几何思想的启发.

17.2.2 罗巴切夫斯基和波尔约的非欧几何

不管分析得出的结论如何，陶里纳斯确信他的有关虚球面的几何不适用于任何"实"情形. 那些公式只是非实情形下的一些美妙的结论. 但是由于沙克里和朗贝特在尝试中均没成功推翻锐角假设，其他数学家开始相信，应存在一个平面几何使该假设正确，并且此几何应以陶里纳斯公式为分析基础. 最早是俄国的罗巴切夫斯基(Nikolai Ivanovich Lobachevsky, 1792—1856) 和匈牙利的波尔约(János Bolyai, 1802—1860) 以足够的信心公开发表了他们的新观点. 二人都从平行公设问题入手，打算找出推翻锐角假设的正确方法. 并且两人逐渐都改变了观点.

1826 年，罗巴切夫斯基给喀山大学作了一次演讲，扼要地介绍了一种几何，这种几何具有过给定点有多于一条的直线平行于给定直线的性质. 三年后，他用俄文出版了《关于几何的基本原理》，延伸了演讲中的观点. 在接下来的十多年里，他发表了一系列论文，阐述了对新几何的研究，包括 1840 年以德文出版的一篇详细的摘要《平行线理论的几何研究》. 波尔约也是在 19 世纪 20 年代开始他绝大多数创造性的研究工作. 1831 年. 将他的论文(拉丁文)《绝对空间的科学》以附录的形式附于其父 F. 波尔约(Farkas Bolyai) 的一本几何著作中，他的观点独立于未曾被认为是先验的公理 XI[平行公设] 的正确与否. 因为罗巴切夫斯基和波尔约的思想被证明非常相似，我们将以《几何研究》作为基础，集中于前者的研究.



图 17.5 俄罗斯邮票上的罗巴切夫斯基.

人物小传	尼古拉 · 伊万诺维奇 · 罗巴切夫斯基(1792—1856)(Nikolai Ivanovich Lobachevski)
	罗巴切夫斯基(图 17.5) 出生于莫斯科以东 250 英里的俄国城市诺夫哥罗德(现在的高尔基)，具有波兰血统. 14 岁时，他获得进入喀山大学的奖学金，条件是获得学位后，至少需教书六年. 事实上，他整个的成年生涯都在这所大学里度过. 1816 年他成为副教授，1822 年，30 岁的时候荣升为正教授. 在他大学时代，高斯的朋友 M. 巴特尔斯(Martin Bartels) 曾应邀来喀山这所新的大学讲授数学，这影响他转向了数学. 罗巴切夫斯基不仅成为喀山的一名能干的教师，而且还担任许多管理职位，包括图书馆馆长和校长. 在法国革命和随之而来的俄国白色统治下，他仍在力所能及的范围内努力将教职工聚集在一起. 不幸的是，最终他还是在 1846 年被无理地剥夺了职位，潦倒地度过了最后的 10 年.

最初，罗巴切夫斯基给出一个确实不依赖平行公设的特定的几何结论的摘要. 接着，他明确给出平行的新定义："相对于平面上一给定直线而言，此平面上通过一点的所有直线可分为两类 —— **相交和不相交**. 这两类线的**边界线**被称为**平行**于所给直线."[5] 这样，如果 BC 是一条直线，A 为不在此线上一点，AD 是到 BC 的垂线，首先引一条线 AE 垂直于 AD(图 17.6). AE 线不与 BC 相交. 罗巴切夫斯基接着假定可能有其它线通过 A，例如 AG，不管它延伸到多远，也不与 BC 相交. 经一条交线，

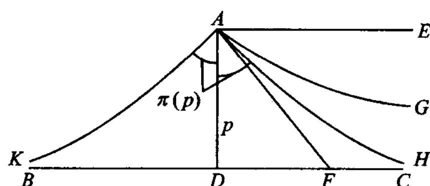


图 17.6　罗巴切夫斯基的平行角.

例如AF,到一条不相交线,例如AG,必然存在一条线AH是两个集合的分界.AH平行于BC.AH和垂线AD间夹角HAD依赖于AD的长度p,此角即是罗巴切夫斯基所称的**平行角**,写作$\Pi(p)$.如果$\Pi(p)=90°$,则经过A只有一条线平行于BC,即为欧几里得平行公理.但如果$\Pi(p)<90°$,AD的相对AH的另一侧,必将存在对应的线AK与AD得到同样的角$\Pi(p)$.在这种非欧几里得的情况下,区分两条不同的平行边常常是必不可少的.无论如何,在非欧假设下,在AD的每一侧,有无穷多条经过A的直线不与BC相交.

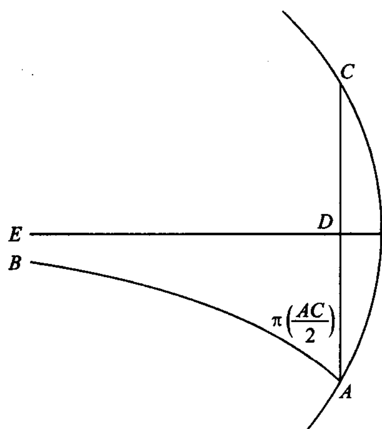


图 17.7　罗巴切夫斯基的极限圆.

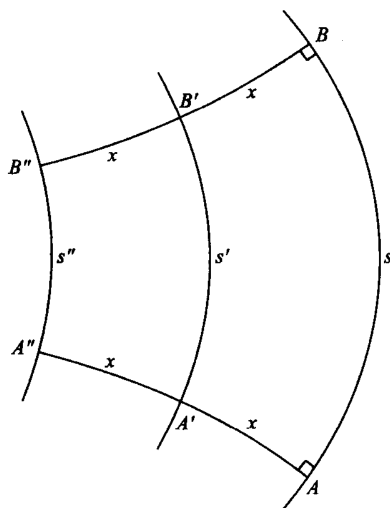


图 17.8　极限圆的垂线.

从非欧假设中,罗巴切夫斯基得到许多结论,其中有些从本质上说是沙克里或朗贝特早已知晓的.例如,他指出,对任意p,$\Pi(p)<90°$这个性质等价于任一三角形之和小于180°,在这种情况下,不仅对任一小于直角的角来说,方程$\Pi(p)=\alpha$可解,而且平行线彼此渐近.为了更精确地界定平行线的特性,罗巴切夫斯基定义了一个新曲线:"我们所说的**界线**(或**极限圆**)为一平面曲线,过其弦中点的所有垂线彼此平行."[6] 换言之,给定一线AB,A在极限圆上,若任意其他点C使AC和AB夹角为$\Pi(AC/2)$,则C为极限圆上的点,因此在此情形下,AC过中点的垂线DE平行于AB(图 17.7).事实上,可证明所有垂直于极限圆的垂线是平行的.我们可以把这条曲线看作无限大半径的圆周,这个圆周在欧几里得假设下为一条直线.设A,B为极限圆上距离为s的两点,引垂直于极限圆的两条垂线AA',BB',$AA'=BB'=x$,罗巴切夫斯基经A',B'建立一个新的极限圆,设$A'B'$的距离等于s'(图 17.8).接着他指出s'/s只依赖于距离x,即$s'/s=f(x)$.如果用相同的方式建造一个新的极限圆$A''B''$,它到$A'B'$的距离为x,且$A''B''=s''$,则同样证明出$s''/s'=f(x)$,因而$f(2x)=s''/s=f(x)^2$.类似地,$f(nx)=f(x)^n$,并且罗巴切夫斯基断定对某一常数a,$s'=sa^{-x}$.因度量单位任

意,他取 a 为 e.平行线 AA', BB' 间的距离由函数 $s' = se^{-x}$ 给出,这里 x 是从极限圆上 A 和(或) B 外引测出的距离.从而得出平行线确实是渐近的.

非欧几里得平面的三角学,是罗巴切夫斯基得到的最有趣的结论,而这些结论沙克里和朗贝特未曾得到.通过对球面三角形和非欧平面上的三角形的研究,罗巴切夫斯基可以以

$$\tan\frac{1}{2}\Pi(x) = e^{-x}$$

的形式明确地计算出函数 $\Pi(x)$,且所获得的结论实质上与陶里纳斯的相同.可得 $\Pi(0) = \pi/2$(或对很小的 x 值,此几何接近欧几里得几何)及 $\lim_{x\to\infty}\Pi(x) = 0$.罗巴切夫斯基得出任意非欧三角形中三边 a,b,c 和对顶角 A,B,C 间新的关系:

$$\sin A\cot\Pi(b) = \sin B\cot\Pi(a), \tag{17.3}$$

$$\cos A\cos\Pi(b)\cos\Pi(c) + \frac{\sin\Pi(b)\sin\Pi(c)}{\sin\Pi(a)} = 1, \tag{17.4}$$

$$\cot A\sin C\sin\Pi(b) + \cos C = \frac{\cos\Pi(b)}{\cos\Pi(a)}, \tag{17.5}$$

$$\cos A + \cos B\cos C = \frac{\sin B\sin C}{\sin\Pi(a)}. \tag{17.6}$$

当三角形边很小时,罗巴切夫斯基的公式暗含了标准欧几里得公式.从 $\Pi(x)$ 的精确公式可到出:

$$\cot\Pi(x) = \sinh x,\quad \cos\Pi(x) = \tanh x,\quad \sin\Pi(x) = \frac{1}{\cosh x}.$$

由其幂级数展开取次数低于2的项得到的这些双曲函数的近似值给出了

$$\cot\Pi(x) = x,\quad \cos\Pi(x) = x,\quad \sin\Pi(x) = 1 - \frac{1}{2}x^2.$$

将这些近似值代入上述四个公式且忽略次数高于2的项,给出结果:

$$b\sin A = a\sin B,$$
$$a^2 = b^2 + c^2 - 2bc\cos A,$$
$$a\sin(A + C) = b\sin A,$$
$$\cos A + \cos(B + C) = 0.$$

前两个结论是熟知的正弦定律和余弦定律,将后两个和前两个结合后,相当于结论 $A + B + C = \pi$.从而也可得,如果分别用 ia, ib, ic 代替三角形三边 a,b,c,罗巴切夫斯基的结论即转换为球面三角学的标准结论.这样,在虚半径的球面上,罗巴切夫斯基的几何实质上和陶里纳斯的对数球面几何相同.

罗巴切夫斯基很可能从未读过陶里纳斯的著作.所以从他的三角公式中他看到的仅仅是"将[非欧]几何假设视为可能成立的一个充分的根据,因此,"他断言,"除了天文观测,没有任何办法可用来判定普通几何计算的正确性."[7] 高斯也认识到创建一个新的显然正确的,但欧几里得平行公设在其中不成立的几何,表明欧几里得几何不存在"必然性",并且不能不加思索地断言欧几里得几何在我们所生活的世界里成立.通过实验确定物质空间的几何是否是欧几里得几何是必不可少的.事实上,罗巴切夫斯基尝试过一种实验,用恒星的位置作为数据,但结果不得要领.

波尔约也探讨过究竟是欧几里得几何代表"现实"还是非欧几何代表"现实",他认为是不能确定的.事实上,他声称,"确定是否[欧几里得几何]或某一[非欧]几何(以及这个)先验存在的不可能性(脱离假设)留待证明.但是,遗留下来是为了等待一个更合适的契机."[8] 尽管"更合适的契机"

从未出现,波尔约确实获得了和罗巴切夫斯基大体一致的数学结论. 在涉及**绝对几何**方面他更明晰些. 绝对几何由真正独立于平行公设的定理集结而成. 例如,他证明了"在任何直线三角形中,以其边为半径的圆(周长)可被视为边的对顶角的正弦"[9] 在欧氏几何中,这里半径为 r 的圆周长为 $2\pi r$. 此结论即为正弦定律 $a:b:c=\sin A:\sin B:\sin C$,在非欧几何里,这里相应的周长为 $2\pi K\sinh\frac{r}{K}$,对某常数 K(对波尔约而言,一个这样的常数决定一个不同的几何),定理转化为:

$$\sinh\frac{a}{K}:\sinh\frac{b}{K}:\sinh\frac{c}{K}=\sin A:\sin B:\sin C,$$

这个结论相当于罗巴切夫斯基的方程(17.3). 波尔约还证明了一个有趣的结论,在非欧几何中使用欧几里得手段可以构造一个正方形,其面积等于半径为 1 的圆面积. 尽管波尔约和罗巴切夫斯基回答了古老的平行公设问题,但在 19 世纪 60 年代以前,几乎没有得到数学界的响应. 原因很多,包括他们的一些(尽管不是全部) 论文被刊登在不太起眼的地方,而且未使用当时的主要语言. 除此以外,通常将一个全新的观念引入到数学中,总是难以被接受. 可以清楚地看到,很少有人能理解一个非欧平面真正是什么,这是这个匈牙利人和这个俄国人的发现没有立即融入数学主流的原因. 尽管创立者推理正确且逻辑上连贯一致,尽管他们展示出涉及已知函数的显然合理的数学公式,新几何的"现实性" 仍不被接受. 非欧几何直到被视为更一般几何体系的一部分,且建立了这个体系与欧几里得几何的联系后,它才不再被认为是旁门左道.



图 17.9　匈牙利邮票上的波尔约.

人物小传	雅诺斯 · 波尔约(1802—1860)(János Bolyai)
	波尔约(图 17.9) 出生于匈牙利的科洛兹瓦尔(现在的罗马尼亚的克鲁日),在现在称作梯尔古 · 穆瑞斯的地方接受了早期教育,他的父亲 F. 波尔约(Farkas Bolyai)—— 高斯的一位朋友 —— 是那里的教授. 16 岁时,他在维也纳参加了皇家军事学会,成为了一个军官,从此以后服役于提密索拉,阿拉德和利沃夫. 但由于身体状况,他不得不于 1833 年退役. 他的父亲对平行公设问题一直怀有兴趣,多年来就此问题与高斯保持着通信,但没取得任何进展. 最终他厌倦了此事并警告儿子不要涉足这个题目. 尽管如此,波尔约坚持了下来,并于 1823 年告诉他的父亲,在平行理论上,他已经得到了"奇妙的发现". 最终,他于 1831 年公布了这个发现. 尽管波尔约努力发展着他的空间理论,但高斯回复他说,他(高斯) 早已发现了非欧几何的基本思想,这让他极为沮丧,使他放弃了进一步公开自己观点的想法.

17.2.3　黎曼关于几何的假设

第一位创建广义的新几何体系的数学家是黎曼. 在他的题为"关于几何基础的假设" 的就职演说中,他给教授团的教授们陈述了他的观点. 在这次演说中,黎曼很少涉及数学上的细枝末节,但却留出大量的时间解释几何应该成为怎样. 他开始"由一般量的观点尝试建立多重延量的概念. 一个多重延量应该可以容纳不同的度量关系,因此空间构成的仅仅是三重延量的一个特殊情况. 但是,由此必然可得出一个结论,几何定理不能由一般量概念推出,而那些可将空间从其他可想象的三重延量中区分开来的性质,只能由经验得出. "[10] 换言之,对黎曼而言,最重要的一般几何概念就是今

天所谓的流形.在一个流形上,人们可以建立不同的度量关系及不同的确定距离的方式.(三维)欧几里得几何,一般被认为是我们所生活的物质空间,其所涉及的是通常的"空间",即是三维流形的特殊情形,欧几里得度量在无限小的意义下可表达为 $ds^2 = dx^2 + dy^2 + dz^2$.像高斯一样,他也主张物质空间的精确性质不是先验地决定,而只能由"经验"决定.在演说的第一部分,黎曼谈到 n 维流形的观点.他归纳地建立了这样的一个流形,由一维流形或曲线开始,"其本质特性是,从流形上任一点起,只可能在两个方向上进行连续运动,向前和向后."[11] 通过使一个一维流形连续通过成为另一个一维流形,创造出一个二维流形;这个二维流形即是由所有通过这种传递所形成的点组成.类似地,一个二维流形连续的传递到另一个二维流形形成一个三维流形,依次类推,可得到更高维的流形.黎曼的中心思想显然是在一点处引入一个更高维数,将进一步导致增加一个可移动的方向,或者用更现代的术语,增加一个新维数即增加了流形在一点上的切空间.黎曼进一步指出,若当函数的零点被定义在 n 维流形上时,从某种意义上说,上述过程可被颠倒过来并定义出一个 $(n-1)$ 维流形.我们可以将这些函数取为今天所谓的**坐标函数**.从而流形上的每一点由 n 个量,即 n 个坐标决定.

黎曼演说的第二部分论述了有关流形度量关系的思想,及一种不依赖于其位置即决定流形上一条曲线的长度的方法.这是仅有的涉及到数学公式的一部分,但即使在这一部分,黎曼也是未加任何推导地直接提出结论.以高斯早些著作为基础的黎曼的基本假设是将曲线上无穷小距离元素 ds 表述为 dx_i 的齐次正定二次函数的平方,即

$$ds^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij}dx_i dx_j,$$

这里 $g_{ij} = g_{ji}$,且任一 g_{ij} 为流形上的连续函数.普通(欧氏)空间是具有这种度量的最简单的情况,即 $ds^2 = \sum dx_i^2$.黎曼指出,在一般情况下,不能将所给度量转化为另一个度量.因此,具有这种最简单度量的流形构成一种特殊的类,一种黎曼称之为"平坦"的类.

为了处理弯曲的(或非平坦的)流形,黎曼构造了一套特殊的坐标,即今天所说的黎曼法坐标,通过使用此坐标,黎曼定义了由高斯思想推广而来的曲率的概念.同时他也指出,曲率对流形是内蕴的并且只依赖于系数 g_{ij}.流形的所有的几何特性均可由度量和坐标网格描述.例如,可以使用渐近平行线和相关的极限圆作为二维曲面的坐标网格,并且使用一适当的度量,在不作对给定直线存在平行线的任何假设下,即可得到罗巴切夫斯基非欧几何的所有性质.黎曼指出常曲率流形具有一重要的性质"图形可在其中无伸缩地移动".事实上,"既然流形的度量性质完全由曲率决定,因此,任一点周围所有方向和任一其他 … 方向完全相同,所以,在常曲率流形中,图形可置于任意位置上."[12]

黎曼演说的最后一部分谈到他的思想和我们通常的三维(欧氏)空间概念的关系.黎曼给出了三组条件,其每一组条件都足以确定一个三维流形是否是平坦的.一组条件是所有实体自由移动和旋转且所有三角形具有相同的内角之和.黎曼指出,确定这种条件是否存在是有些困难的,因为必须将我们的观察推广到无限大或无限小.但他确实指出能够假定物质空间构成了一个三维无界流形.但无界并不意味着空间无限,因为如果曲率是正恒定的,空间必然有限.至于无限小,黎曼断言道,度量关系不一定与一般情况下一致.而且,事实上,只要空间每一可测部分的总曲率接近零,曲率是随点而变的.但是,精确地确定曲率及相关度量,是物理学而不是数学的事.

17.2.4 赫姆霍尔兹和克利福德的几何体系



图 17.10 德国邮票上的赫姆霍尔兹.

黎曼从未试图出版他的演说稿,原因可能是他最初没打算作有关几何的演说,并且那时他正忙于其他项目.这样,尽管演说给高斯留下了极深的印象,它的新思想几乎没在其它场合发生作用,直到黎曼英年早逝,才于 1868 年被出版出来.但是一经发现,黎曼的著作即获得了广泛称赞.特别是德国的赫姆霍尔兹(Hermann von Helmholtz,1821—1894)(图 17.10) 和英格兰的克利福德(William Clifford,1845—1879),二人深受黎曼著作影响并发表了他们对此的解释和扩充,这有助于引起更广泛群体的注意.

赫姆霍尔兹在黎曼演讲稿出版后不久,即在一篇论文上以和黎曼明显相似的题目,"论几何基础的一些事实",试图列出对任何合理的几何研究的基础提供一组假设.首先,和黎曼一样,他假定 n 维空间为一个流形.他的定义比黎曼的稍明确的是,他假定点的附近有 n 个独立坐标存在,且当此点移动时至少有一个坐标随之连续变化.从他的例子中,可以清楚地知道他并没有要求相同的一组坐标总是适用于整个流形.赫姆霍尔兹的第二个公理是断言刚体存在.这个假设使通过重叠,判断两个不同的空间的物体相等成为可能.第三,赫姆霍尔兹断言刚体可以自由移动.换言之,对这样的物体上的任一点可被移到空间的任一其他点处;物体上其他点将随着这次移动被移到其他点处,而这些点的坐标经一组特殊的方程与第一个点的坐标相联.进一步假设 $n = 3$,使赫姆霍尔兹得出他自己的关于我们所生活的物质空间的概念,即一个三维常曲率流形的概念.由此得出对物质空间的三种可能性:曲率为正,负或零.第三种选择导致欧氏几何.相反地,"如果曲率测度为正,我们得到的是**球面**空间,其中直线首尾相接且无平行线.这样一个空间,就像一个球的表面,是无限的但不是无限大.另一方面,一个负恒定曲率测度给出**伪球面空间**,此空间中直线延伸至无限远,并可以通过任何点在任一平坦曲面上画出一束直线,在此面上它将不与另一条给定直线相交."[13] 这样,赫姆霍尔兹成功的将罗巴切夫斯基的非欧几何置于黎曼假设的前提下.而且,球面几何同样被证明为一种平行线不存在的非欧几何.这样,两种否定欧几里得平行公设的理论均导致我们的物质空间的可能几何学.

19 世纪 70 年代早期,克利福德在英格兰进行了一系列演讲,也试图确定物质空间的公设.比赫姆霍尔兹更为明确的是,他指出,通过相似公设,有一种方法来区分欧氏空间和非欧空间,"不改变其形状,任何图形可以以任何程度放大或缩小."[14] 可以证明,这个公设相当于零曲率假设,因而对罗巴切夫斯基的几何来说是不正确的.但是,罗巴切夫斯基思想所造成的革命给予克利海德特别深刻的印象,他把这些思想对于欧几里德作家的作用比作为哥白尼的思想对托勒密体系的作用一样.在这两个情形中,人类对宇宙的看法都有了根本的改变.特别地,以正或负曲率的非欧几何表示的物质世界的可能形式证明了人类对那个空间的知识,特别在那遥不可及的地方的知识被局限在观测能力所能达到的程度之内了.

1876 年,克利福德在一篇短文中也指出,尽管限定于实验精度之内,空间的有限部分确实具有零曲率,但事实上对于空间非常小的部分,我们并不真正知道所有的空间公理是否可以适用.事实上,克利福德提出了一些与 赫姆霍尔兹的关于物质空间具有常曲率的概念相矛盾的新推测.为解释这些在更近时代成为宇宙论研究前沿的新思想,他写道,

事实上,我认为

1. 空间的小部分有一种类似于曲面上的小山的性质，而这曲面平均来说是平坦的．即，几何的普通定律对这些小的部分是不成立的．
2. 呈弯曲的或扭变的这种性质以波浪方式连续的从空间的一部分传到另一部分．
3. 空间曲率的这种变化，确实如我们称之为**物质运动**的那种现象中所发生的情况一样，不管这种物质是有重量的还是像空气一般稀薄．
4. 在这个物理世界中，只有（可能地）遵循连续性规律的这种变化，而无其他．[15]

克利福德的关于物理世界的假设使黎曼的流形理论的思想成为重要的物理研究工具．事实上，20 世纪早期，在物理学有关相对论的变革性的发展中，这些思想被证明是处于中心地位．

17.2.5　非欧几何模型

因为罗巴切夫斯基的几何在常负曲率曲面上似乎正确，曾在博洛尼亚、比萨和罗马数学界享有声誉的意大利数学家贝尔特拉米（Eugenio Beltrami，1835—1900），尝试建立所谓的**伪球面**．结果是，在欧几里得三维空间仅可以建立曲面的一部分．尽管如此，在这个曲面上，贝尔特拉米成功地确定了合适的度量，并给出了非欧空间下这种度量和罗巴切夫斯基的三角定律的关系．对某一值 a，他首先将位于欧几里得三维空间的半径为 k（曲率为 $1/k^2$）的球面参数化：

$$x=\frac{uk}{\sqrt{a^2+u^2+v^2}},\quad y=\frac{vk}{\sqrt{a^2+u^2+v^2}},\quad z=\frac{ak}{\sqrt{a^2+u^2+v^2}}.$$

接着通过代入欧几里得形式 $ds^2=dx^2+dy^2+dz^2$，直接计算球面上度量形式 ds^2：

$$ds^2=k^2\frac{(a^2+v^2)du^2-2uv\,du\,dv+(a^2+u^2)dv^2}{(a^2+u^2+v^2)^2}.$$

为将此公式变形为曲率为 $-1/k^2$ 的伪球面上的结论，贝尔特拉米用 iu 替换 u，用 iv 替换 v．这样所得到的度量

$$ds^2=k^2\frac{(a^2-v^2)du^2+2uv\,du\,dv+(a^2-u^2)dv^2}{(a^2-u^2-v^2)^2}$$

就具有所需要的性质．[16]

在伪球面上，对任一常数 $c<a$，曲线 $u=c$ 和 $v=c$ 是测地线，且分别正交于 $v=0$ 和 $u=0$．这样，贝尔特拉米考虑了一个顶点在原点的直角三角形，一条侧边沿着曲线 $v=0$，一条沿曲线 $u=c$，斜边沿过原点的一条测地线，且和 $v=0$ 的夹角为 θ（图 17.11）．经过对适当的度量形式积分，他计算出三边长度．对斜边，令 $u=r\cos\theta$ 和 $v=r\sin\theta$．从而，

$$ds=\frac{ka\,dr}{a^2-r^2}.$$

对此弧元，容易从 0 到 r 积分，从而可得斜边长度 ρ：

$$\rho=\frac{1}{2}k\ln\frac{a+r}{a-r}.$$

类似地，沿曲线 $v=0$，$ds=\dfrac{a\,du}{a^2-u^2}$．三角形这个侧边长 s 对于给定 u 为

$$s=\frac{1}{2}k\ln\frac{a+u}{a-u}=\frac{1}{2}k\ln\frac{a+r\cos\theta}{a-r\cos\theta}.$$

最后，给出 $u=c$ 上度量

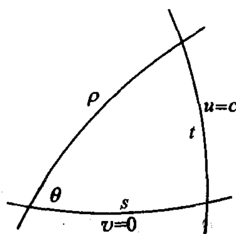


图 17.11　贝尔特拉米在伪球面上的计算：

$$\rho=\frac{1}{2}k\ln\frac{a+r}{a-r},$$

$$s=\frac{1}{2}k\ln\frac{a+r\cos\theta}{a-r\cos\theta},$$

$$t=\frac{1}{2}k\ln\frac{\sqrt{a^2-u^2}+v}{\sqrt{a^2-u^2}-v}.$$

$$ds = \frac{k\sqrt{a^2 - u^2}}{a^2 - u^2 - v^2}dv,$$

此微分的积分适用于一个特殊的 v：

$$t = \frac{1}{2}k\ln\frac{\sqrt{a^2 - u^2} + v}{\sqrt{a^2 - u^2} - v}.$$

对值 ρ，s 和 t 进行代数操作后，贝尔特拉米给出

$$\frac{r}{a} = \tanh\frac{\rho}{k} \quad \frac{r}{a}\cos\theta = \tanh\frac{s}{k} \quad \frac{v}{\sqrt{a^2 - u^2}} = \tanh\frac{t}{k}.$$

从而有

$$\cosh\frac{s}{k}\cosh\frac{t}{k} = \cosh\frac{\rho}{k}.$$

这个结果，完全等同于直角三角形情况下的陶里纳斯的方程（等式 17.1）和罗巴切夫斯基的方程（等式 17.4），从而表明贝尔特拉米的曲面及相关的度量给出了和罗巴切夫斯基的非欧几何相同的几何.换言之，贝尔特拉米的计算表明陶里纳斯看似神秘地使用虚半径球面相当于在合适的二维流形中引入一个新度量.

用一种不同的方法来看罗巴切夫斯基的几何，是仅考虑这个虚球面被投影在圆 $u^2 + v^2 = a^2$ 的内部，这里 u 和 v 是上述给定的参数.可以证明，此圆（图 17.12）上的弦表示罗巴切夫斯基平面上的直线.平行线是相交于圆周上的线，并且，圆周表示“无限”处的点.圆内部不相交的弦表示在罗巴切夫斯基平面根本不相交的线.贝尔特拉米未曾明确计算出此模型中点点间的距离，不过，1872 年，这个缺口被克莱因填平，他使用了一些将在 17.3.3 节讨论的射影几何的概念.1882 年，庞加莱（Henri Poincaré，1854—1912）发展了类似的圆内部的罗巴切夫斯基几何模型.在这个模型中，正交于边界圆的圆弧表示直线.相交于边界的圆弧表示平行线.此模型的优势在于用欧几里得方法可测出圆间角.图 17.12 说明为什么三角形角度之和小于 π.

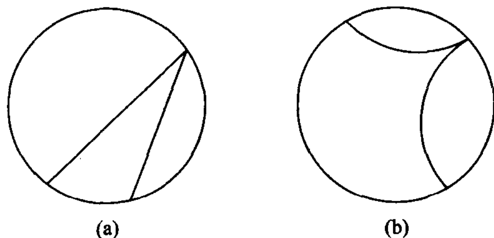


图 17.12　(a) 罗巴切夫斯基平面上克莱因的平行直线；(b) 罗巴切夫斯基平面上庞加莱的平行直线.

19 世纪末，正是使用作为普通欧几里得平面子集的罗巴切夫斯基几何模型，才有助于使数学家们确信非欧几何和欧氏几何同样有效.前一个几何中任何矛盾经变换至模型，导致了后者的矛盾.沙克里为“维护”欧几里得的努力失败了，而罗巴切夫斯基、波尔约、贝尔特拉米、克莱因和庞加莱的著作，却无疑可以明白地证明欧几里得是对的.早在 2200 年前，他将平行公设作为公设的这个结论完全正确.因为罗巴切夫斯基代替欧几里得平行公设的公设导致了和欧几里得同样有效的几何，从而将公设作为一个定理去证明是不可能的.

17.3　射影几何

18 世纪晚期蒙日（Monge）在画法几何上的成就，特别是通过不同类型的射影将三维物体投影

在二维平面上，在19世纪早期再次引起了对射影几何的形式研究，即几何图形的射影不变性研究的兴趣．早在文艺复新时期，射影几何的某些方面，即作为艺术家们努力掌握透视理论的一部分而得到研究．17世纪中期德扎格和帕斯卡开始了这个学科理论的研究，但直到19世纪数学家们才扩大了这项研究的领域．

17.3.1 彭赛列和对偶性

彭赛列(Jean-Victor Poncelet, 1788—1867)是蒙日的学生，他于1822年发表了综合射影几何学领域的第一篇论文《图形的射影性质》．彭赛列是从圆锥曲线的极线理论入手的．给定圆锥曲线 C，任一点 P 可相配一条**极线** π，π 是连接由 P 引向 C 的切线的切点的直线．类似地，对任何与圆锥曲线相交的直线 π'，可相配一点 P' 为其**极点**，极点是过 π' 与 C 的交点的圆锥曲线两条切线的交点．彭赛列看出这些概念是对偶的，即如果 P' 在 π 上，则 P' 的极线 π' 经过 π 的极点 P(图17.13)．

除了极线和极点的对偶性，彭赛列建立起了点和线间更一般的对偶的概念．他注意到通常情况下，涉及"点"和"线"的一个为真命题，如果把这两个词交换一下，命题还是正确的．例如，语句"过两个不同的点恰好决定一条直线"成为"过两条不同的直线恰好决定一个点"(注意后一语句在普通的欧几里得几何中不正确，但是当无穷远点被当作平面上平行线的交点时，后者是正确的．这个观点下面将会讨论到)．给一个更复杂的例子，回忆一下帕斯卡定理："若一个六边形的六个顶点位于一圆锥曲线上，则三组对边的交点位于一条直线上"．其对偶定理是："若一六边形的六条边相切于一条圆锥曲线，则连接三对对顶点的直线相交于一点．"尽管彭赛列没有将对偶原理确立为定理，但他确实将它作为科学发现的有用工具来使用．

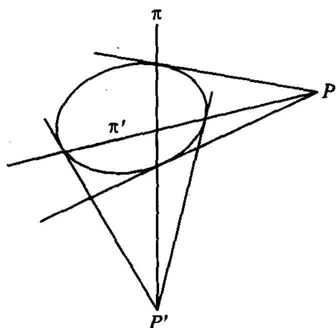


图17.13 极点和极线的对偶关系．

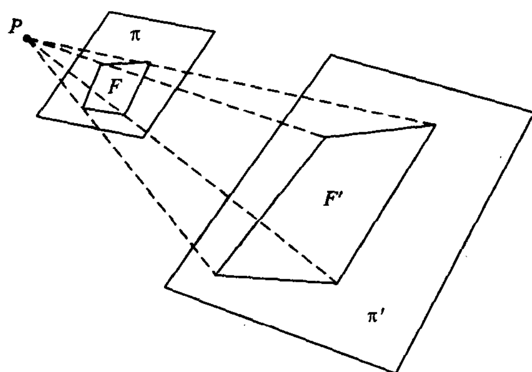


图17.14 F 在 F' 上的中心投影．

但是，彭赛列发现的结果和其论文的首要目标是中心投影的性质．给定平面 π 上的图形 F 和平面外的一点 P，F 对另一平面 π' 的中心投影是过 P 点经 F 上的点得到的所有直线与 π' 的交点所组成的图形(图17.14)．例如，π 上一个正方形在 π' 上的投影是一个四边形，不一定是正方形，而圆周的投影是一个圆锥曲线．彭赛列的目的是确定在这样的投影下，图形的哪些性质是恒定不变的．显然，线段的长度不是投影不变量，但一条直线和一个圆周最多相交于两点这个性质就是这样的一个不变量．彭赛列注意到因为投影可以将平行线变换为相交线，若投影保持相交性不变，则必须引入用来充当一般平行线交点的无穷远点．那么，假定所给平面上所有无穷远点组成无穷远的一条直线

是有用的.这样彭赛列的想法引出了对一个新的对象即射影平面的研究,这个射影平面由平面上一般的点和无穷远点组成.但是,在射影平面图形中,无法辨别一个一般(欧几里得)点和一个无穷远点,因为中心投影总可将一般点变成无穷远点,反之亦然.

为了处理无穷远点,在射影平面上建立一个坐标系被证实是必不可少的.1831 年,普吕克(Julius Plücker,1801—1868)引入了齐次坐标,从而做到了这一点.平面上直角坐标为(X,Y)的一点 P,若 $x = Xt$,且 $y = Yt$,齐次坐标为(x,y,t).在这个定义下,一个点不只具有惟一一组坐标,两组坐标因相差一个常数倍而不同.尽管如此,这些坐标的使用意味着任何多项式方程 $f(X,Y) = 0$(在直角坐标下)被改写成形如 $g(x,y,t) = 0$.这里 g 的所有项具有相同的次数(这即是"齐次").此外,射影平面的无穷远点具有齐次坐标$(x,y,0)$.普吕克注意到,射影平面上的所有直线在这些坐标下,具有方程 $ax + by + ct = 0$.即,给定常量(a,b,c),符合这个方程的点集$\{(x,y,t)\}$都在一条特殊的线上.但使人惊讶的是,如果我们将(x,y,t)看作常量,这个方程也表述出所有经过给定点(x,y,t)的线集$\{(a,b,c)\}$.这样,彭赛列在对偶原则中交换"点"和"线",由在方程 $ax + by + ct = 0$ 中交换"常量"和"变量"而在代数意义下被证明是对的.

17.3.2　交比

彭赛列没有发现最重要的射影不变量——一条直线上四点的交比.但这个概念以非调和比的名字由他的年轻的同胞沙勒(Michel Chasles,1793—1880)得到充分的研究.回想如果直线 p 上的线段 AB 经由 S 点的中心投影投向直线 p' 上的线段 $A'B'$,则 $A'B'$ 的长度一般来说不同于 AB 的长度.类似地,如果 C 为 AB 上一点,C' 为 $A'B'$ 上的对应点,比例 $AC : CB$ 不同于比例 $A'C' : C'B'$.但若 C 和 D 是线段 AB 上的两点,其在线段 $A'B'$ 上的对应点为 C',D',则分别地,交比

$$\frac{AC}{CB} : \frac{AD}{DB}$$

在投影下不变,即:

$$\frac{AC}{CB} : \frac{AD}{DB} = \frac{A'C'}{C'B'} : \frac{A'D'}{D'B'}.$$

这个结果不难证明.引线段 A_1B 和 AB_2 平行于直线 p' 并确定 C,D 在这两条线上的投影(图 17.15).从相似三角形 ACC_2 和 BCC_1,及相似三角形 ADD_2 和 BDD_1,可以看出

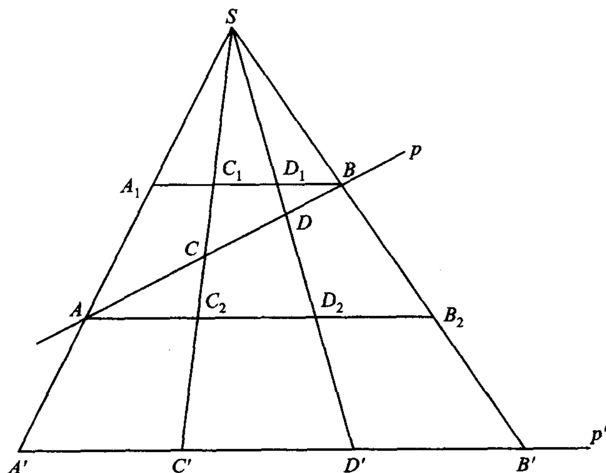


图 17.15　在中心投影下交比不变.

$$\frac{AC}{CB}=\frac{AC_2}{C_1B} \quad 和 \quad \frac{AD}{DB}=\frac{AD_2}{D_1B}.$$

但 $D_1B/C_1B = D_2B_2/C_2B_2$,从而

$$\frac{AC}{CB}:\frac{AD}{DB}=\frac{AC_2}{C_2B_2}:\frac{AD_2}{D_2B_2},$$

即,直线 p 上的交比等于由线段 AB_2 决定的直线上的交比.由基本的相似原理可以容易地得出,后一线段上点的交比等于直线 p' 上的交比.

在前面段落中的交比的标准记法为 (AB, CD).经对四个字母进行排列,在这四个点间可以计算 24 个交比.(在这个情况下,一条线段,例如 AB,如果 A 在 B 的左边,则线段 AB 被认为正,反之则为负.)沙勒注意到在这 24 个交比中,真正不同的只有 6 个,且甚至这 6 个也紧密相关.例如,$(AB, CD) = 1/(AB, DC)$ 及 $(AB, CD) = 1-(AC, BD)$.

有趣的是,尽管射影几何旨在研究不依赖长度这样的概念的图形的性质,事实上,交比定义的真正的基础就是线段的长度,是施陶特(Christian von Staudt,1798—1867)在 1847 年纠正了这个问题,他为射影几何概括出一个公理系统,这个系统以保留调和四元组的射影映射的概念为基础.一个**调和四元组**是四点集 A, B, C, D,其 $(AB, CD) = -1$.尽管这个定义好像仍需要长度,施陶特指出给定三个共线点 A, B, C,可以找到"第四个调和"D,即通过简单的射影构造,使 $(AB, CD) = -1$ 的点.在射影几何中,施陶特的工作是非常重要的,为这门学科给出了明确的研究范围,并且为在无度量几何中定义距离的概念提供了思想.

17.3.3 射影度量和非欧几何

凯莱于 1859 年首次为射影平面提出度量的定义.给定一个圆锥曲线 C,他给出一个相当复杂的函数 $d_C(P_1, P_2)$ 的定义,此定义依赖于 C,且满足基本的距离性质,即,若 P_1, P_2, P_3 位于同一直线上,则 $d(P_1, P_2) + d(P_2, P_3) = d(P_1, P_3)$.12 年后,克莱因注意到如果凯莱的圆锥曲线是一个圆周,则在圆周内部的射影平面部分可以被认为是罗巴切夫斯基几何的一个模型,并且对那个几何来说,凯莱的度量可转化为一个距离函数.

根据交比,克莱因为非欧几何平面定义了修正的度量.考虑在圆周 $u^2 + v^2 = 1$ 内的罗巴切夫斯基平面.在这个圆周内部给定两点 P, Q,用直线连接它们,直线与圆周相交于点 R, S(图 17.16),P 到 Q 的直接距离为

$$d(P,Q) = c\ln(QP, RS) = c\ln\left(\frac{QR}{RP}:\frac{QS}{SP}\right) = c\ln\left(\frac{QR\cdot SP}{RP\cdot QS}\right),$$

其中 c 为某一个决定长度单元的常量.直接可以证明,如果 P, Q, Q' 是这条直线上的三点,则 $d(PQ) + d(QQ') = d(PQ')$.所以函数 d 满足一个距离函数的主要特性.例如,置 $P = R$,同样可以表明整个弦的长度为无穷,即圆周本身代表了无穷远点.克莱因在罗巴切夫斯基平面上的度量相当于贝尔特拉米在伪球面上所获得的度量.为说明之,取 P 为开始点,Q 为 P 右边欧几里得距离 $r < 1$ 的一点.则

$$d(P,Q) = c\ln\left(\frac{-(1+r)\cdot(-1)}{1\cdot(1-r)}\right) = c\ln\left(\frac{1+r}{1-r}\right).$$

这样,合适地选择 c,可得到贝尔特拉米曾计算的相同的距离值.

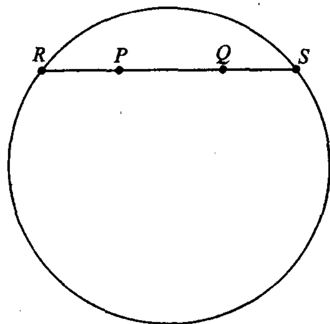


图 17.16 克莱因的罗巴切夫斯基距离:$d(P,Q) = c\ln\left(\frac{QR}{RP}:\frac{QS}{SP}\right)$.

克莱因使用交比也给出平面上两直线间角的定义.但此定义以一个略微不同的形式更易理解些.假定借由三元组(a_1,b_1,c_1),(a_2,b_2,c_2)以射影坐标给定两条直线,即经方程$a_1x+b_1y+c_1t=0$和$a_2x+b_2y+c_2t=0$,这些线交于点$x_0=b_1c_2-c_1b_2, y_0=c_1a_2-a_1c_2, t_0=a_1b_2-b_1a_2$.它们的相交角$\alpha$则为

$$\alpha = \arcsin \frac{\sqrt{t_0^2 - x_0^2 - y_0^2}}{\sqrt{(a_1^2 + b_1^2 - c_1^2)(a_{22} + b_2^2 - c_2^2)}}.$$

注意,此公式暗含着若直线交点在边界圆周$x^2+y^2=t^2$,即$(x/t)^2+(y/t)^2=1$或$u^2+v^2=1$上时,相交角为0.

克莱因进一步指出,随着对边界曲线的不同的选取,一个类似的距离定义要么导致欧氏几何,要么导致非欧球面几何.(克莱因称球面几何为**椭圆几何**;称欧氏几何为**抛物几何**;称罗巴切夫斯基几何为**双曲几何**.)事实上,克莱因修改了球面非欧几何,因为,当它存在时,两点并不总是决定惟一一条直线.修改之处为应将球面上直径方向相对的点看为相同的.新"半球面"将不能在普通三维空间中存在(因为沿赤道相对的点认为是相同点),但可以通过乘一个虚常量一特定交比的对数来定义一距离函数.于是可以证明测地线的总长为πR(这里R是球面半径).克莱因通过使用相同类型的距离函数将欧氏几何和两种非欧形式的几何统一起来,这是使数学家确信非欧几何和欧几里得几何形式是一致的另一个因素.

17.3.4 克莱因的爱尔兰根纲领

1872年,克莱因在他的爱尔兰根纲领中给几何研究作出了另一重要贡献.这篇论文详述了这样一个想法,通过将一般几何视为:研究在基本空间(或流形)中,在一特殊变换群作用下保持不变的图形特性,则19世纪各种几何研究均可统一起来并可对之归类.事实上,克莱因展示的几何和变换群之间的关系,有助于在19世纪末期,为群的抽象观念的发展提供了原动力.

克莱因的入手点可能是他意识到保留了他的罗巴切夫斯基平面模型中的边界圆周不变的射影平面到自身的射影变换保持交比不变,因而距离和角不变.这些变换在罗巴切夫斯基平面模型上是刚体运动.研究了一些早期的群理论著作,克莱因进一步认识到因为在这个集合中,任两个变换的合成属于这个集合且这个集合任何变换的逆变换也属于它,所以保留边界曲线的所有射影变换集合组成了一个变换群.而且,因为在这个群作用下,这个平面上图形的基本特性保持不变,人们可将这个非欧平面几何明确视为对这些不变特性的研究.这样,1872年,克莱因定义了一般"几何"为:"给定一个流形和流形的变换群;来研究流形的关于在变换群作用下不会改变的特性的构形"[17].

克莱因给出了几个几何和其相伴群的例子.普通的二维欧氏几何对应于克莱因所谓的**主群**,主群是由伴有相似变换和反射所组成群的平面上的所有刚体运动.在这些变换下的不变量形成了经典欧氏几何所研究的对象.射影几何由在投影下,或等效的是,在直射变换即把直线变到直线的变换下保持不变的图形的研究组成.这些变换可以以

$$x' = \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}, \quad y' = \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}$$

的形式解析地表示出来,这里$\det(a_{ij}) \neq 0$.因为可将主群解析的表示为形如

$$x' = ax - by + c, \quad y' = bex + aey + d$$

的变换集,其中$a^2+b^2 \neq 0$.且$e=\pm 1$.显然,这是射影群的一个子群.从而后者的不变量少于前者,因此射影几何的任何定理在欧氏几何中仍然成立,但反之不成立.

人物小传	费利克斯·克莱因(1849—1925)(Felix Klein) 起初,克莱因在哥廷根大学负责创建数学学会,这个学会在20世纪的三分之一时间中,将一所毫不起眼的大学改变为世界的数学中心.在其早期光辉的生涯中,他在几何研究上成绩斐然,但后来在19世纪80年代中期,他精神上受到一次重创,此后,他将精力倾注于教学,写书和组织活动.他成为 Mathematische Annalen——那个时代最好的数学期刊之一的——的编辑,并将自己针对不同主题所作的演讲编辑成册,许多书专门写给数学教师并涉及到高中数学教育的核心问题.此外,他的主要著作还有19世纪数学史及指导《数学科学的百科全书》的出版,这本百科全书旨在收集到那时为止的可获得的所有数学成果和方法,但这个目的最终也未能达成.

克莱因的几乎没有详细分析的《爱尔兰根纲领》,目的在于成为几何研究的基本纲领.但直到19世纪90年代被翻译为意大利文,法文和英文后,它才受到关注.从那以后,克莱因的变换群不变量的思想在几何的任何领域都成为研究的重要对象,并在进入20世纪后成为几何研究的一个中心.

17.4　n 维几何

我们已经看到在19世纪初期,许多数学家是如何突破希腊几何限制于三维的例子.例如,19世纪30年代,奥斯特罗格拉茨基在几乎很偶然的情况下(仅在不同公式后增加三个圆点),将他的发散量定理推广到 n 维,而更早些,柯西在解释他的对称矩阵对角化的观点中,即论述了任意维的几何对象.实际上,术语"n 维几何"似乎最早出现在1843年的凯莱的论文题目中,而论文本身是纯代数的,涉及到几何的部分只是一带而过.

17.4.1　格拉斯曼和《扩张论》

第一位系统提出维数大于3的空间理论的是 H.格拉斯曼(Hermann Grassmann,1809—1877),格拉斯曼是德国的数学家和语言学家.不幸的是,在他的有生之年,其光辉著作未能得到承认.在1844年的《线性扩张论》和1862年的修订本中,格拉斯曼从几何乘法的概念开始,建立了一套用符号来表达几何思想的系统方法.

他第一次系统写出 n 维空间著作的四年之前,在一篇关于潮汐理论的新解释的论文中,格拉斯曼已能处理二维和三维空间中向量的乘法.他定义**二个向量的几何积**为"由这些向量决定的平行四边形的面容度"及**三个向量几何积**为"由它们所形成的立体(平行六面体)."[18] 通过合适的途径给这样的积定义符号,他能说明,二个向量几何积满足分配律和反交换律,而且位于同一平面的三个向量其几何积为0.因为二个向量几何积即平行四边形的面积等于二向量长度及其夹角正弦值的乘积,此积在数值上恒等于现代的叉积.当然,差异存在于乘出后产生的结果其几何性质不同.这个积不是作为一个新向量,而是作为一个二维对象.但由于在三维空间,格拉斯曼的平行四边形和现代的向量一一对应(若两平行四边形面积相同且位于相同或平行的平面上则认为其相等)此一一对应由每一平行四边形的相伴法矢决定,此向量长度等于平行四边形的面积,所以两种乘法本质上是一样的.但格拉斯曼积不像叉积一样,其优势在于可以推广到高维.推广正是格拉斯曼1844年和1862年主要著作的基础.

在他的文章中，特别是在 1862 年那个更清楚的版本中，格拉斯曼从具有固定长度和方向的线段这样的向量概念开始他的探讨. 二个向量的相加是通过将第二个向量的起点连到第一个向量的终点这种标准化的方式得到. 减法只是加上负向量，即具有相同长度和相反方向的向量. 向量是格拉斯曼所谓的**扩张量**的最简单的例子. 通常，这样一个量被抽象地定义为“用单位体系中的数导出的任何表达式”[19]. 这里，格拉斯曼指的是先从线性无关量的集合$\epsilon_1, \epsilon_2, \cdots \epsilon_n$ 开始，再取这些“单位”的任何线性组合作为一个扩张量. 扩张量经由简明的方法相加：$\Sigma\alpha_i\epsilon_i + \Sigma\beta_i\epsilon_i = \Sigma(\alpha_i + \beta_i)\epsilon_i$. 类似的，可以由一个数量去乘扩张量. 格拉斯曼指明了适用于扩张量的代数基本定律，并定义了量$\{\epsilon_i\}$的**空间**为它们所有线性组合的集合.

接下来，格拉斯曼用分配律$(\Sigma\alpha_i\epsilon_i)(\Sigma\beta_j\epsilon_j) = \Sigma\alpha_i\beta_j[\epsilon_i\epsilon_j]$，这里每一个$[\epsilon_i\epsilon_j]$称作二阶量，定义了扩张量的乘法. 因为这个新和式必是一个扩张量，它也一定可以表示为单位元的线性组合. 这样，格拉斯曼需要定义二阶单位. 他证明出，假定将定义在单位上的乘法法则扩充为对任何扩张量都适用的法则，则对定义二阶单位而言，只有四种基本可能. 第一，所有$[\epsilon_i\epsilon_j]$无关. 第二，可以有$[\epsilon_i\epsilon_j] = [\epsilon_j\epsilon_i]$（这个乘法满足所有一般代数乘法法则）. 第三，可以有$[\epsilon_i\epsilon_j] = -[\epsilon_j\epsilon_i]$（这暗含着对所有 i，$[\epsilon_i\epsilon_i] = 0$）. 最后，所有的积$[\epsilon_i\epsilon_j] = 0$. 这个积的第三种形式，叫做**组合**乘法，这是格拉斯曼在他其余的著作中主要处理的课题. 由他的条件，对任何一阶扩展量 A 和 B，乘法法则$[AB] = -[BA]$成立.

定义了两个一阶单元的组合积后，格拉斯曼使用相同的法则，直接定义了三个或三个以上的一阶单位的积. 例如，如果有三个一阶单位$\epsilon_1, \epsilon_2, \epsilon_3$，则存在三个二阶单位$[\epsilon_1\epsilon_2], [\epsilon_2\epsilon_3], [\epsilon_3\epsilon_1]$和一个三阶单位$[\epsilon_1\epsilon_2\epsilon_3]$（其他三个一阶单位的任何积具有两个相同的因子，因而等于 0）. 如果有四个一阶单位，则有六个二阶单位，四个三阶单位，一个四阶单位. 格拉斯曼进一步指出，n 个一阶单位的 n 次线性组合的积$(\Sigma\alpha_{1i}\epsilon_i)(\Sigma\alpha_{2i}\epsilon_i)\cdots(\Sigma\alpha_{ni}\epsilon_i)$等于$\det(\alpha_{ij})[\epsilon_1\epsilon_2\cdots\epsilon_n]$，这里，括号的表达式是个单个的 n 阶单位.

人物小传	赫尔曼·格拉斯曼（1809—1877）（Hermann Grassmann）
	格拉斯曼出生并大半生居住在波兰的波美拉尼亚州，现在波兰什切青的斯德丁市. 尽管在柏林大学他主要学习语言学和神学，但离开大学后，他却回到斯德丁从事数学和物理工作以准备通过这些科目的教师州级考试. 后来他曾短期任教于一所柏林技术学校，并在 1836 年后，供职于家乡的一些学校. 他一生中最大的抱负即是在大学里取得职位，但尽管他发展了向量空间理论，很少有人看到他的努力或认识到他的伟大的创造力. 格拉斯曼将他的书的副本寄给一些有影响力的数学家，但只有一位，黎曼的学生汉克尔（Hermann Hankel）对它称赞有加，并计划将格拉斯曼的一些材料收入他自己有关复变量的书中. 19 世纪 60 年代，格拉斯曼将他的注意力转向语言研究，并对梵文研究作出了许多重要的学术贡献. 但他后来的数学著作质量不是很高，他也没能达到成为一名大学教授的目标.

格拉斯曼在《扩张论》中所确定的组合积，用现代术语，即为向量空间的外代数. 它的思想来源于格拉斯曼想到符号化地表达不同的几何概念. 所以，特别的，他将二阶量想作平行四边形，三阶量想作平行六面体. 尽管没有对更高阶量做特殊的几何解释，格拉斯曼看出符号操作不需要限制于任何维数上. 他不仅构造了外代数，而且，他发展了许多有关向量空间即有关 n 单位所有线性组合空间的重要思想. 早在 1840 年，他就已建立两个了向量内积的概念为“一向量乘以由第二个向量对它的垂直投影的代数积”[20]. 并指明用坐标的形式，内积可由$(\Sigma\alpha_i\epsilon_i)(\Sigma\beta_j\epsilon_j) = \Sigma\alpha_i\beta_i$给出. 在《扩张论》

里，他给出线性无关和基的概念，指出任何向量可被惟一地表达为基元素的线性组合，并证明出了在 n 维空间中，任何基向量可以由另一个与剩余 $n-1$ 个向量无关的向量代替.他证明了一个量的正交系是线性无关的(这里若内积为 0 则称两向量正交)，他还证明了一个著名的结果，即对一个空间 V 的两个子空间 U, W，有

$$\dim(U+W)=\dim U+\dim W-\dim(U\cap W).$$

格拉斯曼的努力终其一生都未受到重视，但 19 世纪晚期，他的线性代数和外代数的思想重新被发现，并被用于数学的许多新领域，包括微分形式理论和向量空间理论.

17.4.2　向量空间

线性代数的基本概念，包括线性无关和线性组合，在 19 世纪就被用于数学的一些领域.但直到 19 世纪末，向量空间的抽象定义才形成.1888 年，佩亚诺在他的《几何演算》首次给出这样一个定义.正如题目所指出的那样，和格拉斯曼一样，佩亚诺此书的目的即是为了发展几何计算.这样，此书的很大一部分篇幅花在各种点，线，面和立体图形的计算上.但在第 Ⅸ 章，佩亚诺给出他所说的线性系统的定义.这样一个系统，由具有加法和数乘运算的量组成.加法必须满足交换律和结合律(尽管这些定律不是由佩亚诺命名的)，而数乘满足两个分配律，一个结合律和对每个量 v，有定律 $1v=v$.此外，佩亚诺将满足对任何 v，$v+0=v$ 和 $v+(-1)v=0$ 的零量的存在作为公理体系的一部分包含在内.佩亚诺也定义了线性系统的**维数**为系统中线形无关量的极大数目.与这种想法有关的是，佩亚诺指出，一个变量的多项式函数组成的集合形成一个线性系统，但线性无关量的这样一个极大数是不存在的，因此，这个系统的维数必然无穷.

像格拉斯曼一样，佩亚诺的著作没有立即在数学界产生影响，尽管数学家们继续使用着他所论及的基本概念，但他的定义却被遗忘了.例如，1893 年，戴德金在其有关代数数域的著作中，定义了一个空间 Ω，Ω 是 n 个无关的代数数的，系数在一个域中的所有线性组合集合所组成的空间.他注意到，这个空间的数满足一些我们认为是属于向量空间的基本性质，而他并未在任何地方提及一个这样的定义.使用归纳法，他证明出一个重要结论，在 Ω 中，任意 $n+1$ 个向量相关.尽管他没有明确指出没有更小的生成元集合决定这个空间，但他的定义本质上已假定了这个结论，这就说明(有限维)向量空间的维数已被很好的定义出了.

在数学文献中，有关向量空间的理论层出不穷，但一直到 20 世纪，对这个主题的完全的公理化处理才进入数学主流.

17.4.3　微分形式

格拉斯曼外积的最重要的应用之一是由埃利·嘉当(Elie Cartan，1869—1951)发展的微分形式的理论.微分形式，自然是“积分号下的东西”，在 19 世纪得到广泛的使用，特别是线积分，面积分和体积积分.但是，没有人尝试去定义形式本身，定义的只是积分.嘉当读过格拉斯曼的著作后，于 19 世纪 90 年代后期，提出在 n 维空间里，取单位系统为微分 $dx_1, dx_2, dx_3, \cdots$，从而可以定义微分形式.这些单位的乘法是格拉斯曼的组合积，而单位的系数将为空间的微分函数.这样，二维下的一形式是形如 $A(x,y)dx+B(x,y)dy$ 的表达式，三维下的二形式具有形式 $A(x,y,z)dxdy+B(x,y,z)dydz+C(x,y,z)dzdx$，乘法遵循法则 $dx_idx_j=-dx_jdx_i$，因此 $dx_idx_i=0$.

当然，嘉当认识到这个组合乘法将解决欧拉的如何找到决定变量变换公式的形式化方法的问题.因为若函数 $u=u(x,y)$，$v=v(x,y)$，定义变换从 x, y 到 u, v，则

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy, \quad dv = \frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy,$$

且积 $dudv$ 由

$$\begin{aligned} dudv &= \left(\frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy\right)\left(\frac{\partial v}{\partial x}dx + \frac{\partial v}{\partial y}dy\right) \\ &= \frac{\partial u}{\partial x}\frac{\partial v}{\partial x}dxdx + \frac{\partial u}{\partial x}\frac{\partial v}{\partial y}dxdy + \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}dydx + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y}dydy \\ &= \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right)dxdy \\ &= \frac{\partial(u,v)}{\partial(x,y)}dxdy \end{aligned}$$

给定,这和期望的一样.

除了发展微分形式的代数,嘉当也发展了它们的微积分.即,1899 年,他定义了一形式 $\omega = \Sigma A_i dx_i$ 的导数表达式(现在所说的外导数)为二形式 $d\omega = \Sigma dA_i dx_i$.例如,形式 $Adx + Bdy$ 的导数表达式为形式

$$\begin{aligned} d\omega &= \left(\frac{\partial A}{\partial x}dx + \frac{\partial A}{\partial y}dy\right)dx + \left(\frac{\partial B}{\partial x}dx + \frac{\partial B}{\partial y}dy\right)dy \\ &= \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\right)dxdy. \end{aligned}$$

注意这个导数表达式出现在格林定理的表述中,而一形式 $Adx + Bdy + Cdz$ 的外导数出现在斯托克斯定理的表述中.在 1901 年,嘉当将他的外导数定义推广到了任意阶的形式中.即,如果 $\omega = \Sigma a_{ij\cdots k}dx_i dx_j \cdots dx_k$,则外导数 $d\omega$ 定义为 $\Sigma da_{ij\cdots k}dx_i dx_j \cdots dx_k$.于是可直接给出二形式 $Adydz + Bdzdx + Cdxdy$ 的外导数为三形式

$$\left(\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z}\right)dxdydz,$$

此表达式出现在散量定理中.

尽管嘉当意识到这三个向量微积分定理用微分形式可以很容易的表述出来,但直到 1917 年,才由古尔萨(Edouard Goursat)(1858—1936)首次提出这些定理的沃尔泰拉推广,即今天所谓的广义斯托克斯定理,可以用简单形式改写为

$$\int_S \omega = \int_T d\omega,$$

这里 ω 是 n- 空间的 p- 形式,且 S 为 $(p+1)$ 维区域 T 的 p 维边界.古尔萨也使用微分形式陈述并证明了庞加莱引理和它的逆定理,即如果 ω 是 p- 形式,则 $d\omega = 0$ 当且仅当有 $(p-1)$- 形式 η,使得 $\omega = d\eta$.但古尔萨没有注意到,结果的"仅当"部分依赖于 ω 的域并非总是正确.1922 年,嘉当本人给出了一个反例,这成为 20 世纪 30 年代微分流形的微分上同调得以发展的推动力之一.[22]

17.5　几何基础

19 世纪晚期,出现了为各种类型的数学结构构造的公理体系.群和域的概念,就像向量空间的概念那样被公理化了.同样地,为正整数集合建立了公理,并且人们在实数构想的准确定义上作出了极大的努力.当然,存在的最古老的公理体系是欧几里德的几何研究的公理体系.事实上,此体系为这个时期创立各种公理体系提供了模型.但欧氏体系存在许多缺陷.特别是,那个时代许多数学

家都注意到在一些证明中，欧几里得所作的假设在他的公理和公设体系中没有被明确的提出. 随着非欧几何的新发展，数学家们重新检验了各种公理的特性，这样，不足为奇的是，许多数学家协同努力来修改欧几里得著作，从而使欧氏几何拥有尽可能牢固的根基.

17.5.1 希尔伯特公理

在试图建立完整的，可导出欧氏几何的公理集合的数学家中，最成功的大概要推 19 世纪末 20 世纪初的数学家希尔伯特(David Hilbert,1862—1943). 在 1899 年，希尔伯特出版了他的《几何基础》，此书基本上收录了他在 1898—1899 年在哥廷根大学秋季学期所作的关于几何的演讲. 这本著作旨在"以一种尽可能清楚的方式，将不同公理集的意义和个别公理所导出的结论范围阐释出来，为几何学选择一个简单而完备的互相独立的公理集，并由之推导出最重要的几何定理."[23]

希尔伯特由三个未定义的术语，点、直线和面着手，并借助公理定义了它们的相互关系. 正如希尔伯特指出的那样，只有公理才能定义它们的相互关系. 在证明结论中不应该使用几何直觉. 事实上，可以简单地用其它概念来替代这三个概念 —— 希尔伯特建议用椅子，桌子和啤酒杯，只要它们满足公理. 因此，可以看出，希尔伯特的公理体系思想和欧几里得，亚里士多德的稍微有点不同. 希腊思想家们仅试图论述某种关于他们已经直觉上理解的概念的"显然的" 事实. 另一方面，希尔伯特像那些论述群公理的人那样，决定不再对所期望的性质作任何具体的解释. 这样，任何物体都能成为一"点" 或一"线"，只要"点" 和"线" 满足几何公理，就如同只要满足群公理的物体满足"乘法" 法则一样，任何物体集合都可以成为一个群.

希尔伯特将他的公理分成五组：关于连接的，关于次序的，关于平行的，关于全等的，关于连续性和完全性的. 第一组包括七个公理，建立了三个基本概念：点，线和面之间的联系. 这样，两点不仅决定一条直线(公理 I,1)，而且仅决定一条直线(公理 I,2). 类似的，不在同一直线上的三点决定一个且只决定一个平面(公理 I,4). 第五个公理断言，如果直线上的两点位于一给定平面上，则整条直线也位于这个平面上. 第六个公理讲，任两个有一个公共点的平面至少有另一个公共点. 最后，第一组第七个公理断言，每一条直线上至少存在两点，每一个平面上至少存在三个不共线点，一个空间至少存在四个不共面点. 列出这些公理后，作为结论，希尔伯特指出，一个平面上的两条直线要么有一个公共点要么没有，两个平面要么没有公共点要么有一条公共直线.

人物小传

大卫 · 希尔伯特(1862—1943)(David Hilbert)

希尔伯特是距我们最近的世界级的顶尖数学家之一，在数学的许多领域他都作出了突出的贡献，他生命的前 33 年在过去的东普鲁士的首都，现在属于俄罗斯的哥尼斯堡(Königsberg)及其附近度过. 在那里他上了大学并获得了博士学位，这之后，于 1885 年成为大学教员，但只在 F. 克莱因邀请他到哥廷根大学后，他才被授以重任，不久他便成为使哥廷根大学在 20 世纪前三分之一的时间里，接替柏林大学，成为德国乃至世界上数学的优秀学府的主要功臣之一. 希尔伯特的学术生涯开始于研究代数形式，接着转向代数数论，几何基础，积分方程，理论物理及最后研究的数学基础. 他最为人所津津乐道的大概要数 1900 年，在巴黎召开的国际数学家大会上所作的演讲. 在那里，他提交了一份写有 23 个他认为将成为 20 世纪数学的重要问题的清单. 希尔伯特坚信，正是问题驱使数学进步，同样，一直以来他始终确信"wir mussen wissen, wir werden wissen(我们必须知道，我们将会知道)." 纳粹夺权后，希尔伯特被迫目睹了他所熟知并深爱着的哥廷根大学分崩离析，在二次世界大战中，他孤独的死去.

希尔伯特的第二组公理使他将线段 AB 定义为位于 AB 两点间的点集.欧几里得本身没有明确假定“中间性”的特性,很可能由于含此特性的图表的“显然性”.希尔伯特通过公理化“中间性”的思想,明确了欧几里得的假设.例如,公理 II,3 断言,直线上任意三点总是有且仅有一点介于另两点中间,而公理 II,5 断言,经三角形一边上一点,且不经过任一顶点的任一直线必经过另两边中某一边上的一点.由这些公理,希尔伯特能推导出一个重要的定理:任一简单多边形可将平面分为两个不相交区域,内部和外部,且连接内部任一点和外部任一点的直线必和多边形有一个公共点.

第三组公理仅由希尔伯特的平行公理组成:“在一平面 α 上,在直线 a 外,过任一点 A 可以引出一条且仅有一条直线和直线 a 不相交.”[24] 第四组假定了基本的全等思想的存在.当时,欧几里得是通过将一个三角形“放置”到另一个三角形上来证明他的第一个三角形全等定理.许多人对这种叠放方法的正确性提出了置疑.出于这个原因,希尔伯特对全等这个未定义的术语列出了六个公理.例如,公理 IV,1 断言,给定线段 AB 和一点 A',存在这样的一点 B',使线段 AB 全等于线段 $A'B'$.而公理 IV,2 实质上确定了全等是一种等价关系.类似的,将角定义为从一点发出的两条不同的半射线所组成的系统后,希尔伯特断言,给定一个角和一条射线,可以确定全等于第一个角的另一个角.本组的最后一个公理几乎肯定了《原本》的 I4:如果两个三角形有两边及其夹角全等,则剩余的角也全相等.希尔伯特没有断言第三边全等,但通过矛盾论证明了它.在这些公理的帮助下,他也能够证明另两个三角形全等定理,欧几里得的所有直角彼此相等的公设,内错角定理和三角形内角之和等于两直角定理.

希尔伯特最后一组公理由两个公理组成,涉及了连续性的基本思想.首先,是阿基米德公理:假定 A,B,C,D 为四个不同的点,则在射线 AB 上存在不同点 $A_1,A_2,\cdots,A_n$ 的有限集,这样的每一线段 A_iA_{i+1} 全等于线段 CD 且使得 B 介于 A 和 A_n 间.换言之,给定任一线段和任一测度,存在一整数 n,这样的 n 个测度单元生成一条大于给定线段的线段.此公理若与先前的公理结合使用时,可得到线的长度无限的结论.这样,这个欧几里得的不言而喻的假设,在沙克里和兰伯特否决钝角假设中起了重要作用,现在得到了明确.希尔伯特最后一个公理本质上论述了直线上的点和实数一一对应.换言之,直线上没有“洞”.这个公理回答了对欧几里得在《原本》I,1 中构造等边三角形所持的异议,即不能保证所构造的两个圆确实相交.根据希尔伯特公理,没有点可以被添加到这两个圆上,因此它们不能仅仅是从一个“穿过”另一个.

17.5.2 相容性、独立性和完备性

论述完这些公理,希尔伯特进而指出它们都是**相容的**,即至少在算术不矛盾的假定下,从其中不会推导出任何矛盾.他的观点在表明非欧几何没有矛盾上和克莱因及其他人类似.即仅使用满足公理的算术运算,就可以构造一个几何.例如,由一个特定的代数数集 Ω 开始,希尔伯特定义点 p 为 Ω 中一个有序数对(a,b),并定义直线 L 为 Ω 中三个数的比$(u:v:w)$,这里 u,v 不能全为 0.则若 $ua+vb+w=0$ 时,p 在 L 上.希尔伯特以用这种方式算术地阐明了的每个几何概念以及在此阐述中所满足的每个公理,为几何创建了一个公理的算术模型.如果公理在几何中导致矛盾,则在算术中将出现类似的矛盾.因此,假如算术的公理相容,则几何的公理也相容.

一个公理系统的另一个重要特征是**独立性**,即,没有一个公理是从其余的公理中推导而来.尽管希尔伯特没有完整地证明独立性,但他确实指出了,通过建立几个有趣的模型,某一组公理满足它,而其它的不满足,可以说明不同组公理独立.例如,他建构了一个除阿基米德公理外,其他公理均满足的系统.希尔伯特没有论述到公理系统深层次的特性 —— **完备性**.即可以指出,任一可在系统中表述的语句是真是假.但事实上可以肯定,希尔伯特相信他的系统是完备的.实际上,不久就有

几个数学家指出,欧氏几何的所有定理都可以用希尔伯特公理证明.

希尔伯特的成就不仅在于他驳回了各种对欧几里得推理系统的反对意见,而且加深了这样一个观念:任何数学领域都必须始于未定义的术语并明确说明术语间的关系.如上所述,19世纪晚期,许多公理系统被发展来阐明不同的数学领域.希尔伯特的工作被认为是这个过程的顶峰.因为他能够选取这个最古老的系统,并说明经过一点修补它可以经得住时间的考验.这样,在19世纪末,欧几里得和亚里士多德的数学思想作为纯数学模型,再一次得到证实.一个世纪之后,这些思想仍然盛行.

习　题

高斯微分几何中的问题

1. 证明顶点为(x,y),$(x+dx,y+dy)$,$(x+dx,y+\delta y)$的(无穷小)三角形的面积为$\frac{1}{2}(dx\delta y-dy\delta x)$.

2. 如果一个曲面由方程$z=z(x,y)$给出,试证明曲率k可表示为

$$k=\frac{z_{xx}z_{yy}-z_{xy}^2}{(1+z_x^2+z_y^2)^2}.$$

提示:首先证明如果X,Y,Z是单位球面上的坐标,而点$(x,y,z(x,y))$在曲面上,则有

$$X=\frac{-z_x}{\sqrt{1+z_x^2+z_y^2}},\quad Y=\frac{-z_y}{\sqrt{1+z_x^2+z_y^2}},\quad Z=\frac{1}{\sqrt{1+z_x^2+z_y^2}}.$$

3. 计算抛物面$z=x^2+y^2$的曲率函数.

4. 假设$x=x(u,v)$,$y=y(u,v)$,$z=z(u,v)$是某曲面的参数方程,并且$E=x_u^2+y_u^2+z_u^2$,$F=x_ux_v+y_uy_v+z_uz_v$,$G=x_v^2+y_v^2+z_v^2$,证明

$$dx^2+dy^2+dz^2=Edu^2+2Fdudv+Gdv^2.$$

5. 设单位球面的参数方程为$x=\cos u\cos v$,$y=\cos u\sin v$,$z=\sin u$,试计算上题中出现的E,F,G,并且证明$ds^2=du^2+\cos^2 u dv^2$.

非欧几何中的问题

6. 设单位球上有一个球面三角形,其三边为a,b,c,对应的角为A,B,C,试通过把此三角形分成两个直角三角形,并应用第4章的公式来证明如下公式:

$$\cos a=\cos b\cos c+\sin b\sin c\cos A.$$

7. 证明在第6题中,如果球半径为K,则上面的公式可写为

$$\cos\frac{a}{K}=\cos\frac{b}{K}\cos\frac{c}{K}+\sin\frac{b}{K}\sin\frac{c}{K}\cos A.$$

8. 利用幂级数,证明当$K\to\infty$时,陶里纳斯的"对数－球面"公式

$$\cosh\frac{a}{K}=\cosh\frac{b}{K}\cosh\frac{c}{K}-\sinh\frac{b}{K}\sinh\frac{c}{K}\cos A.$$

9. 对虚半径为i的球面上的渐近直角三角形,即$\sin B=\frac{1}{\cosh x}$,证明陶里纳斯公式等价于平行线角,即$\tan\frac{B}{2}=e^{-x}$的罗巴切夫斯基公式.

10. 证明虚半径为$\mathrm{i}K$的球面上的半径为r的圆的周长是$2\pi K\sinh\frac{r}{K}$,而且当$K\to\infty$时,这个值趋于$2\pi r$.(提示:首先确定半径为K的普通球面上的圆的周长.)

11. 如果 $\tan \frac{1}{2}\Pi(x) = e^{-x}$，其中 $\Pi(x)$ 是罗巴切夫斯基的平行线角，试导出公式

$$\sin \Pi(x) = \frac{1}{\cosh x} \quad 和 \quad \cos \pi(x) = \tanh x,$$

而且证明它们的二阶幂级数展开式分别是

$$\sin \Pi(x) = 1 - \frac{1}{2}x^2 \quad 和 \quad \cos \Pi(x) = x.$$

12. 当非欧三角形的边 a, b, c 很"小"时，试通过把第 8 题的结果变换成

$$\sin A \tan \Pi(a) = \sin B \tan \Pi(b),$$

$$\cos A \cos \Pi(b) \cos \Pi(c) + \frac{\sin \Pi(b) \sin \Pi(c)}{\sin \Pi(a)} = 1$$

而导出正、余弦定理.

13. 如果 ABC 是边为 a, b, c 的三角形，试从公式 $a\sin(A + C) = b\sin A$、$\cos A + \cos(B + C) = 0$ 及正弦定理导出 $A + B + C = \pi$.

14. 如果三角形的三个边 a, b, c 分别用 ia, ib, ic 代替，则罗巴切夫斯基的基本三角形公式（方程 17.3 – 17.6）就转化为球面三角形的标准公式.（为简单起见，假设角 C 是直角.）

15. 对半径为 k 的由下式给出的球面，试用几何方法描述其 Beltrami 的参数化

$$x = \frac{uk}{\sqrt{a^2 + u^2 + v^2}}, \quad y = \frac{vk}{\sqrt{a^2 + u^2 + v^2}}, \quad z = \frac{ak}{\sqrt{a^2 + u^2 + v^2}}.$$

16. 在练习 15 中，用 iu, iv 分别代替 u, v 则曲率为 $\frac{1}{k^2}$ 的球面就变成曲率为 $-\frac{1}{k^2}$ 的伪球面.

17. 证明：如果伪球面上的直角三角形的边长为 ρ, s, t 则贝尔特拉米的公式就为

$$\frac{r}{a} = \tanh \frac{\rho}{k}, \quad \frac{r}{a}\cos\theta = \tanh \frac{s}{k}, \quad \frac{v}{\sqrt{a^2 - u^2}} = \tanh \frac{t}{k},$$

并且 $\cosh \frac{s}{k} \cosh \frac{t}{k} = \cosh \frac{\rho}{k}$.

射影几何中的问题

18. 试用中心投影把平行线变成相交线.

19. 对交比验证下列关系式：

$$(AB, CD) = 1 - (AC, BD), \quad (AB, CD) = \frac{1}{(AB, DC)}.$$

20. 记交比 (AB, CD)，(AC, DB)，(AD, BC) 分别为 λ, μ, ν，则试证

$$\lambda + \frac{1}{\mu} = \mu + \frac{1}{\nu} = \nu + \frac{1}{\lambda} = -\lambda\mu\nu = 1.$$

21. 证明：如果点 p' 位于点 p 相对于圆锥 C 的极线 π 上，则 p' 的极线 π' 经过 p 点.（提示：先假设 C 是圆周.）

22. 试确定点 $(3,4)$ 和 $(-1,7)$ 的齐次坐标.

23. 试写出直线 $2x - y = 0$ 在无穷远点的齐次坐标.

24. 试确定齐次坐标为 $(3,1,1)$ 和 $(4,-2,2)$ 的点的直角坐标.

25. 试（在直角坐标下）确定经过无穷远点 $(2,1,0)$ 和点 $(6,2,2)$ 的直线的方程.

26. 证明平面上的每个圆周都经过两个无穷远点 $(1,i,0)$ 和 $(1,-i,0)$.

27. 给定三个共线的点 A, B, P，证明如图 17.17 中的 Q 点使得 A, B, P, Q 成为一个调和四元组，即使得交比 (AB, PQ) 为 -1.

28. 利用表示罗巴切夫斯基平面的圆周的内部的距离 d 的克莱因的定义，证明：如果 P, Q, Q' 是直线上的三个点，则 $d(P, Q) + d(Q, Q') = d(P, Q)$.

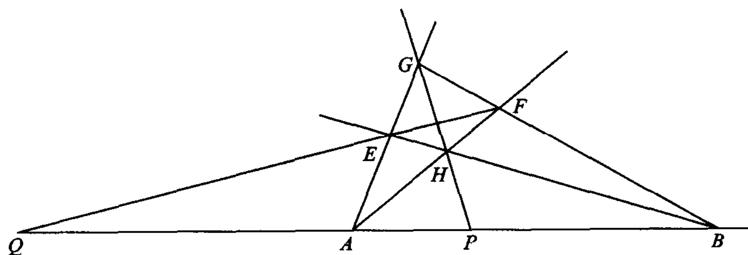


图 17.17 求出 Q，使得$(AB,PQ)=-1$. 经过 A 点引两条直线，经过 P 点引一条直线与这两条直线相交于两点 G 和 H，G 和 H 分别与 B 相连，经过 F 和 E 点（分别是 BG 与 AF，BE 与 AG 的交点）的直线与直线 APB 相交就得到所期望的点 Q.

格拉斯曼领域中的问题

29. 设 i,j,k 是三维空间中的一阶单位元，试确定 $2i+3j-4k$，$3i-j+k$ 和 $i+2j-k$ 的组合乘积.

30. 证明在格拉斯曼的组合乘积中，下列式子成立：

$$\left(\sum a_{1i}\varepsilon_i\right)\left(\sum a_{2i}\varepsilon_i\right)\cdots\left(\sum a_{ni}\varepsilon_i\right)=\det(a_{ij})[\varepsilon_1\varepsilon_2\cdots\varepsilon_n],$$

其中，每个线性组合是 n 个一阶单位元的集合，$[\varepsilon_1\varepsilon_2\cdots\varepsilon_n]$ 上 n 阶单独的单位元.

涉及微分形式的问题

31. 设 $\omega=A\mathrm{d}x+B\mathrm{d}y+C\mathrm{d}z$ 是三维空间中给定的微分 1- 形式，试证

$$\mathrm{d}\omega=\left(\frac{\partial C}{\partial y}-\frac{\partial B}{\partial z}\right)\mathrm{d}y\mathrm{d}z+\left(\frac{\partial A}{\partial z}-\frac{\partial C}{\partial x}\right)\mathrm{d}z\mathrm{d}x+\left(\frac{\partial B}{\partial x}-\frac{\partial A}{\partial y}\right)\mathrm{d}x\mathrm{d}y.$$

32. 试证 $\omega=A\mathrm{d}y\mathrm{d}z+B\mathrm{d}z\mathrm{d}x+C\mathrm{d}x\mathrm{d}y$ 的外导数是

$$\left(\frac{\partial A}{\partial x}+\frac{\partial B}{\partial y}+\frac{\partial C}{\partial z}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

33. 对三维空间中的微分 1- 形式或 2- 形式ω，试证 $\mathrm{d}(\mathrm{d}\omega)=0$.

34. 设 $R^3-\{0\}$ 中的 2- 形式ω 由下式给出：

$$\omega=\frac{x\mathrm{d}y\mathrm{d}z+y\mathrm{d}z\mathrm{d}x+z\mathrm{d}x\mathrm{d}y}{(x^2+y^2+z^2)^{3/2}}$$

证明 $\mathrm{d}\omega=0$，但是不存在 1- 形式η，使得 $\mathrm{d}\eta=\omega$.（提示：如果存在这样的 1- 形式，则由 Stokes 定理，则有$\int_T\omega=\int_T\mathrm{d}\eta=\int_S\eta=0$，其中 T 是单位球面. 因为 T 的边界是空集，所以可直接算出$\int_T\omega$.）

讨论题

35. 研究几种新版的中学几何教材，它们是遵循欧几里得的公理，还是希尔伯特公理体系？或是它们的组合形式？试对在中学几何教学中采用希尔伯特框架的益处给出你的评价.

36. 由陶里纳斯，罗巴切夫斯基和贝尔特拉米所给出的非欧几何的解析形式是否比综合形式要好？怎样解释虚半径球面？

37. 为什么首先在非欧几何方面发表论文的两位数学家都来自 19 世纪的非主流国家？这是偶然的，还是更有深层次的原因？

38. 仔细研读黎曼的讲演“几何基础的假说”. 试描述黎曼的主要新思想，这些思想在 20 世纪又如何得到发扬光大？常常听到这样一种说法：黎曼的工作是爱因斯坦广义相对论的先驱，你对此有何评论？

文献和注解

对 19 世纪的几何历史各个方面的一些重要著作包括有 Jeremy Gray, *Ideas of Space: Euclidean, Non-Eucliden and Relativistic*(Oxford: Clarendon Press, 1989). B. A. Rosenfeld, *A History of Non-Eucliden Geometry. A Critical and Historical Study of Its Development*(New York: Dover, 1955), Julian Lowell Coolidge, *A History of Geometrical Methods*(New York: Dover, 1963), Michael I. Crowe, *A History of Vector Analysis*(New York: Dover, 1985), 另外与群论历史相关的有 Hans Wussing, *The Genesis of the Abstract Group Concept*(Cambridge: MIT Press, 1984).

1. Gauss, 写给 Olbers 的信, 引自 Rosenfeld, *History of Non-Euclidean Geometry*, p. 215.
2. Gauss, *Abstract of the Disquisitiones Generales Girca Superficies Curvas*, 这是篇提交给哥廷根皇家学会的文章, 被收录在 *General Investigations of Curved Surfaces* 中, 由 Adam Hiltebeitel 和 James Morehead 翻译并作注(Hewlett, NY: Raven Press, 1965), p. 45.
3. Gauss, *General Investigations of Curved Surfaces*, p. 10. 这篇文章稍早一些的版本及其历史背景均包括在 Raven Press 的版本中, 颇值得一读. 对那些有微分几何经历的人, 值得尝试把高斯的思想翻译成现代语言. 在 Michael Spivak 的 *A Comprehensive Introduction to Differential Geometry*(Waltham, MA: Brandeis University, 1970), vol. II 中对此有所讨论. D. J. Struik 的文章"Outline of a History of Differential Geometry,"Isis 19(1933), 92 – 120 及 20(1933), 161 – 191 中也讨论了高斯的工作.
4. 同上, p. 21.
5. Lobachevsky, *Geometrical Investigations on the Theory of Parallel Lines* 收在 Bonola 的 *Non-Euclidean Geometry*, p. 13 中, 由 G. B. Halsted 翻译. Bonola 的这本书不仅含有罗巴切夫斯基的这篇基础性文章, 还翻译了波尔约的 *Appendix*. 因此, 对非欧几何的这两种处理方法便能易于进行比较. 关于罗巴切夫斯基更多的信息可见 A. Vucinich, "Nikolai Ivanovich Lobachevski, The Man Behind The First Non-Euclidean Geometry", Isis 53(1962), 465 – 481 及 V. Kagan, *N. Lobachevski and his Contribution to Science*(Moscow: Foreign Language Publishing House, 1957).
6. 同上, p. 30.
7. 同上, pp. 44 – 45.
8. Bolyai, *Appendix exhibiting the absolutely true science of space*, 收在 Bolyai, *Non-Euclidean Geometry*, p. 48 中, 由 G. B. Halsted 翻译.
9. 同上, p. 20.
10. Riemann, "On the Hypotheses which lie at the Foundation of Geometry", 由 Spivak 翻译, 收入他的 *A Comprehensive Introduction to Differential Geometry*, vol, II, pp. 4A – 4 – 4A – 20 中. 除翻译了 Riemann 的演讲外, Spivak 还仔细讨论了这篇著作.
11. 同上, p. 4A – 7.
12. 同上, p. 4A – 15.
13. Helmholtz, "On the Origin and Significance of Geometrical Axioms", 收在 James Newman, *The World of Mathematics*(New York: Simon and Schuster, 1956), vol. 1, 646 – 668, p. 657.
14. Clifford, "The Postulates of the Science of Space", 收在 Newman, *World of Mathematics*, vol. 1, 552 – 567, p. 564.
15. Clifford, "On the Space Theory of Matter", 收在 Newman, *World of Mathematics*, vol. 1, 568 – 569.
16. Beltrami 处理伪球面的度量及它与非欧几何关系的文章被翻为英语并收入 John Stillwell 的 *Sources of Hyperbolic Geometry*(Providence: American Mathematical Society, 1996) 中. 此书也包括了, Felix Klein 和 Henri Poincaré 的基础性文章.
17. Klein, "Vergleichende Betrachtungen über neuere geometrische Forshungen", (Erlangen: Deichert, 1872) 以及 *Mathematische Annalen* 43(1893), 63 – 100. 由 M. W. Haskell 翻译为"A Comparative Review of Recent Researches in Geometry", *Bulletin of the New York Mathematical Society*, 2(1893), 215 – 249, p. 218. 有一些最近的关于克莱因具深远影响工作的文章, 其中有 Thomas Hawkins 的"The Erlanger Programm of Felix Klein: Reflections on its Place in the

History of Mathematics," *Historia Mathematica*, 11(1984), 442 – 470, 以及 David Rowe, "A Forgotten Chapter in the History of Felix Klein's Erlanger Programm," *Historia Mathematica*, 10(1983), 448 – 454.

18. Grassmann, "Theorie der Ebbe und Flut, 引自 Crowe, *History of Vector Analysis*, p.61. Crowe 的著作给出了向量分析历史的方方面面的仔细研究, 特别是讨论了 19 世纪后半叶中四元数的拥护者与向量的拥护者之间的矛盾冲突.

19. Grassmann, *Ausdehnungslehre*(1862), 引自 Eugene Smith, *A Source Book in Mathematics*(New York: Dover, 1959), p.685. 在 J.V. Collins, "An Elementary Exposition of Grassmann's *Ausdehnungslehre*, or Theory of Extension," *American Mathematical Monthly*, 6(1899), 193 – 198, 261 – 266, 297 – 301 和 7(1900), 31 – 35, 163 – 166, 181 – 187, 207 – 214, 253 – 258 中详细讨论了"*Ausdehnungslehre*". 关于格拉斯曼创造线性代数方面的最近的著作是 Desmond Fearnley-Sander 的"Hermann Grassmann and the Creation of Linear Algebra," *American Mathematical Monthly*, 86(1979), 809 – 817. "*Ausdehnungslehre*"的 1844 年版本业已翻成英语, 并与格拉斯曼的其他著作一起, 汇集成一本书: Hermann Grassmann, *A New Branch of Mathematics*, 由 Lloyd C. Kannenberg 翻译(Chicago: Open Court, 1995).

20. Grassmann, "Theorie der Ebbe und Flut," 引自 Crowe, *History of Vector*, p.63.

21. 对向量空间理论和抽象线性代数的起源方面更多的细节见 Jean-Luc Dorier, "A General Outline of the Genesis of Vector Space Theory," *Historia Mathematica*, 22(1995), 227 – 261 以及 Gregorg H. Moore, "The Axiomatization of Linear Algebra: 1875—1940", *Historia Mathematica*, 22(1995), 262 – 303.

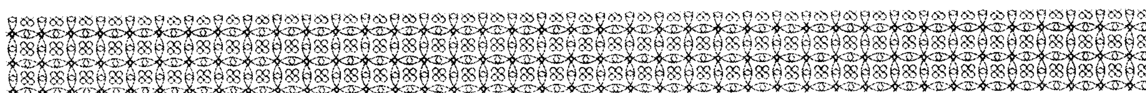
22. 对微分形式发展方面的更多细节见 Victor J. Katz. "Differential Forms – Cartan to DeRham," *Archive for History of Exact Sciences*, 33(1985), 321 – 336.

23. Hilbert, *The Foundations of Geometry*(La Salle, Il: Open Court, 1902), p.1, 也可参见 Michael Toepell, "Origins of David Hilbert's Grundlagen der Geometrie," Archive for History of Exact Sciences, 35(1986), 329 – 344.

24. 同上, p.12.

19 世纪几何学概览

1777—1855	高斯(Carl Friderich Gauss)	微分几何
1788—1867	彭赛列(Jean-Victor Poncelet)	综合射影几何
1792—1856	罗巴切夫斯基(Nikolai Ivanovich Lobachevsky)	非欧几何
1793—1880	沙勒(Michel Chasles)	交比
1794—1874	陶里纳斯(Franz Taurinus)	虚球面上的几何
1798—1867	施陶特(Christian von Staudt)	射影几何
1801—1868	普吕克(Julius Plücker)	对偶、齐次坐标
1802—1860	波尔约(János Bolyai)	非偶几何
1809—1877	格拉斯曼(Hermann Grassmann)	n 维空间几何
1821—1894	赫姆霍尔兹(Herman von Helmholtz)	n 维几何基础
1821—1866	黎曼(Georg Bernhard Riemann)	几何猜想
1835—1900	贝尔特拉米(Eugenio Beltrami)	负曲率曲面
1845—1879	克利福德(William Clifford)	物理空间的公理假设
1849—1925	克莱因(Felix Klein)	爱尔兰根纲领、度量非欧几何
1854—1912	庞加莱(Henri Poincaré)	非欧几何模型
1858—1932	佩亚诺(Giuseppe Peano)	向量空间公理
1858—1936	古尔萨(Edouard Goursat)	广义斯托克斯定理
1862—1943	希尔伯特(David Hilbert)	几何基础
1869—1951	嘉当(Elie Cartan)	微分形式



第18章 20世纪的数学

> 埃米·诺特的思想非常具有启发性.她的许多建议都是在她的学生或合作者的工作中才最后成型……哈塞(Hasse)承认他的关于超复量和类域论之间的联系方面的优美论文应归功于诺特的非正式评论.她能以精确的预言方式发表像"标准剩余符号只不过是循环代数而已"这样有远见的论述,这出自她强大的想象力,而这种想象力绝大多时间中都得到了成功,并随着时光流逝得到了加强;这样的论述从而成为能指明以后困难工作的路标……最重要的是她开创了代数中一种全新的划时代的思想方式.
>
> ——摘自1935年外尔(Hermann Weyl)纪念埃米·诺特的讲话.

阿佩尔(Kenneth Appel)和哈肯(Wolfgang Haken)曾经怀疑是否能完成他们的计划,但1976年7月24日,通过证明1936年的最后的不可避免构形的可约性,他们借助于计算机完成了四色定理的证明.四色定理,即用四种颜色即可为任何地图着色,首次提出于1852年7月26日.这两个人向美国数学会提交了一份报告,该报告刊登在《美国数学会通报》(Bulletin of the AMS)9月刊上.有趣的是在6月中旬哈肯就已经向美国数学会提交了一份他计划在8月份的夏季会议上提交的论文的摘要.该论文题目为"为什么四色问题是困难的?"

随便看一下任何数学方面的研究性图书馆都可以发现20世纪的数学成果远远超过以前各世纪的总和,在那里堆满书架的杂志中,绝大部分开始于这个世纪.甚至在19世纪时是最有力的那些杂志中,比如克雷尔的杂志或刘维尔的杂志,20世纪的部分也占据了它们的大部分篇幅.但是,不管是好还是坏,大学里所教授的大部分内容是从19世纪或更早时候开始的,这样在为大学层次设计的课本中人们对20世纪的数学只能接触到相当浅的表层.因此在这一章里我们将集中于四个精选的20世纪数学领域,这些领域在典型的大学课程中都有一定程度的覆盖.对那些想对20世纪的数学有更深入的探究的读者,图书馆的书架是开放的,那儿有许多作为指导的资源.

我们开始先谈谈在世纪初数学基础方面的一些问题.康托尔在无限集理论的工作在19世纪末刚出现时就立即引起了争议,而且持续到20世纪初.康托尔自己以及其他许多数学家试图证明无限集上的一些看起来"显然"的结果,但收效甚微.比如康托尔尝试证明无限数的三分律,即对任意

两个基数 A,B,恰有以下三种性质之一成立:$A = B, A > B$ 或 $A < B$.他的尝试遇到了未曾预料到的麻烦,就像他在试图证明实数是良序一样.求解这些问题的一个结果是导致了集合论中的一个新的公理 —— 选择公理.这个公理实际上不言而喻地已被使用了许多年,直到 1904 年策梅洛(Ernst Zermelo) 才明确表述出来,但是策梅洛对这个新的公理的叙述引起了新的争议.为解决这个问题,策梅洛给出了集合论的公理化,这个公理化也有助于解决在世纪之交前后罗素(Bertrand Russell) 发现的各种各样的集合论的悖论.策梅洛和其他人都希望一旦有了集合论的一个正确的公理化,算术理论就能基于它之上并且一般的数学能够有一个安全的基础.但是,结果并不像大多数数学家期望的那样,1931 年哥德尔(Kurt Gödel) 建立了他的不完全性定理,宣称任何在其上自然数的算术能够被表示的理论都有不能由这种理论的公理来证明的结果.这样哥德尔的结果从某种意义上来讲终止了数学的公理化进程,即想给出能作为该学科的其他部分的基础的完备和相容的公理系统的尝试.

要讨论的 20 世纪的数学的另一方面是拓扑学的成长,包括点集拓扑和组合拓扑.它们刚刚开始于上个世纪,但注定要成为这个世纪中的数学的一个"生长"领域.点集拓扑根源于康托尔在实数集理论上的工作,它被许多数学家推广到其它许多集合.组合拓扑则根源于黎曼想在有"洞"的区域上求复函数的积分的尝试.但是这样的区域上的理论的开端是庞加莱在 19 世纪 90 年代的工作,特别是他对同调的定义.在 20 世纪早期他的定义通过使用单纯形得到了修改,但是直到 20 世纪 20 年代,组合拓扑和代数的联系才在哥廷根由诺特领导的数学家小组所认识到.

随之而来的拓扑的代数化是我们将考虑的 20 世纪数学的第三方面中的一部分,即在所有数学领域中运用的不断增长的代数技巧.这包括了亨泽尔(Kurt Hensel) 和斯泰尼茨(Ernst Steinitz) 有关域论的一些新思想,特别是巴拿赫(Stefan Banach) 工作中向量空间思想的公理化,以及以诺特的工作为中心的环的新结构的广泛研究.代数领域继续抽象化,其结果导致了艾伦伯格(Samuel Eilenberg) 和麦克莱恩(Saunders Mac Lane) 于 1945 年引进了范畴和函子理论.

代数学在机器计算中发挥了非常重要的作用,这包括第二次世界大战时和稍后的电子计算机的发展.这一章的最后一节将讨论在此进展中起重要作用的数学的许多方面,同时也将介绍计算机与数学实践的一些联系.因此我们将考虑巴贝奇(Charles Babbage) 和拜伦(Ada Byron) 工作中发展计算机的早期尝试,图灵(Alan Turing) 的工作中的可编程计算机的理论基础,以及在冯·诺伊曼(John von Neumann) 领导下的高等研究院的计算机构造的一些方面.我们将以因计算机的发现而产生影响的数学的几个部分的简短介绍作为结束,它们包括纠错码、线性规划和图论.

18.1 集合论:问题与悖论

康托尔在他 19 世纪末的工作中提出了许多关于无限集合理论的问题,他多年尝试解答.大多数情况下未获成功.在进入 20 世纪的前后其他数学家也向这些问题发起进攻.

18.1.1 三分法和良序性

康托尔证明了对两个基数分别为 $\overline{\overline{M}}$ 和 $\overline{\overline{N}}$ 的集合 M 和 N,下列三个关系最多只能有一个成立:$\overline{\overline{M}} = \overline{\overline{N}}, \overline{\overline{M}} < \overline{\overline{N}}$,或者 $\overline{\overline{M}} > \overline{\overline{N}}$.早在 1878 年他就觉得恰有一个关系应该成立,这一点对他来说似乎是显然的.也就是说,如果两个集合没有相同的基数,则其中一个集合必定与另一个集合的一个子集有相同的基数.后来他才意识到否认存在两个集合,使得一个集合均不与另一个集合的一个子集等价并不是一件显而易见的事情.事实上在 1895 年他的《贡献》中他就明确提出了三分法原理,并

强调不仅他没有得到它的证明而且该证明必定是很难的. 因此, 他谨慎地在他的理论的其他证明中避免使用这个原理.

康托尔也意识到这个三分法问题和另一个原理是紧密相关的, 这个原理是说每个集合都是良序的. 也就是说对任意集合 A, 均存在一个次序关系 $<$, 使得 A 的每一个非空子集 B 均包含一个最小元 c 使得对 B 中的其它任意一个元 b 均有 $c<b$. 在自然的次序下自然数是良序的. 另一方面, 在实数的自然的次序下, 实数不是良序的. 但 1883 年康托尔认为一种良序性的存在几乎是不证自明的. 但是到19世纪90年代中期, 他开始觉察到这个结果需要一个证明了. 他在1897年相信找到了证明, 但很快意识到是不完全的.

希尔伯特认为良序原理是如此的重要以致于在为 1900 年于在巴黎召开的世界数学家大会所作的发言中, 他把实数集合是否是良序的问题作为他向数学家们建议在20世纪考虑的23个重要问题中第一问题的一部分提了出来. 如希尔伯特提出这一点时所说:"对我来说最重要的是得到康托尔这个非凡论述的一个直接证明, 或许它可以通过实际给出数的一种排列使在每一个偏序系统之中第一个数都能被指出来证明." 换句话说, 希尔伯特实际希望有人构造出一个自然数的明确的良序.

补遗 18.1　希尔伯特在 1900 年世界数学家大会上的发言

在 1899—1900 年的冬季, 希尔伯特被邀请于 1900 年 8 月在巴黎召开的第二届世界数学家大会上作一个主要发言. 希尔伯特花了好几个月的时间来决定他的发言主题, 最后在七月他才决定在会上将讨论数学上一些有影响的没有解决的问题. 由于新世纪的开始, 他相信这些问题将奠定 20 世纪中数学的基调. 由于希尔伯特的主题确定得太晚, 结果他未能在会议开幕式上发言, 而是在会议的历史和教学的联合会议上发的言, 而这些部分被认为比那些纯数学, 像代数、算术、几何和分析等, 缺乏影响.

希尔伯特以对一个重大问题的标准的讨论来开始他的演说: 它的表述必须清晰易懂, 它应该是困难但不是完全不可着手的, 同时它的解决应该具有重大发展意义. 希尔伯特向他的听众提出的 23 个问题实际上包含了数学的所有分支. 例如, 对数学基础, 希尔伯特要求给出连续统假设的一个证明, 以及对算术公理相容性的研究. 来源于数论的问题是 α^{β} 是否总是代表一个超越数, 甚至恰好是一个无理数?这里的 α 是代数数而 β 是无理数. 若一个丢番图方程是可解的, 这种可能性是否总能够被判定?来自分析的问题是, 是否黎曼函数的所有复零点都有实部 1/2?人们是否总能解决偏微分方程理论中的边界值的问题?

事实上, 希尔伯特的问题已经被证明是 20 世纪数学的中心问题. 许多已经得到了解决, 而剩余的部分都有重大的进展. 或许一位 20 世纪末的数学家会在 1998 年于柏林举行的世界数学家大会上列出一个新的问题清单.

在世纪之初, 集合论中另一个麻烦是出现了许多似是而非的悖论. 最早的悖论之一, 今天称为罗素悖论, 是在 1903 年罗素(图 18.1)(1872—1970) 发表的. 但据策梅洛的一个朋友赫斯尔(Edmund Husserl)(1859—1938) 的一篇笔记所写, 这个涉及到集合包含自身作为元素的悖论似乎是之前一两年前由策梅洛(1871—1953) 发现的:



图 18.1　格林纳达邮票上的罗素.

定理　对于一个集合 M, 如果其所有子集都是它的元素, 则这个集合是不相容的, 即这样一个集合会导出矛盾.

证明:"我们考虑那些不包含它们自身作为元素的子集 $m, \cdots$, 所有这些构成一个集合 $M_0, \cdots$, 现在来证明: (1) M_0 不包含它自身作为元素;

(2)M_0包含自身作为元素.考虑(1):M_0作为M的一个子集,它本身是M的一个元素,但不是M_0的一个元素,否则,M_0将包含以它自身作为一个元素的M的一个子集(即M_0本身)作为一个元素,这与M_0的定义相矛盾.考虑(2):既然M_0是一个不包含它自身作为元素的M的子集.从而它必定是M_0的一个元素."[3]

罗素自己发表了这个悖论的其他几种提法,最简单的是理发师悖论:某个小镇上的一个理发师说他将为并且只为镇上所有那些不能为自己理发的人理发,试问这个理发师要为自己理发吗?

这样到了20世纪早期,康托尔处理集合论的方法尽管在许多新概念的发展上有许多成果,但显然有许多缺陷需要修正,不仅一些貌似显然的结果得不到证明,就连他的一些直观的想法都明显导致矛盾.有趣的是,尽管在这段时期数学的其它许多领域都在进行公理化,康托尔自己并未尝试把他的集合论建立在任何公理体系上.他的定义和证明产生于他的直觉,特别是想想他把集合定义为任何能想到的物体的聚集.正是这种极其宽松的集合定义是产生上述的那些悖论的核心,同时正是缺少合适的公理妨碍了三分法和良序问题的解决.

18.1.2 选择公理

对康托尔的集合论紧密相关的那些异议被策梅洛在1904年到1908年期间扫清了.1904年,策梅洛发表了良序定理的证明,该证明以一个在多年来的各种证明中似乎都隐含出现的原理为基础,即今天所说的选择公理.策梅洛最先对它做了明确的表述:"假设对任何一个非空集合M中的每一个子集M',都有一个任意的元素m'相伴,而它存在于M'自身中;把m'称为M'的特定元".[4]从而,作为一个公理,策梅洛断言总是存在一个"选择"函数,函数$\gamma: S\rightarrow M$(这里S是M的所有子集构成的集合)使得对S中的每一个M',都有:$r(M')\in M'$.换言之,人们总是可以以某种方式从给定的一个集合的每一个子集中"选择"一个元素.

策梅洛意识到它所引入的公理是一个重要原理.它不仅使他能够证明良序定理,而且给了他一个三分法原理的证明.另外他写到:"这个逻辑原理不能 …… 被简化到更简单的形式,但它毫不犹豫地可被应用于数学推导中的每一个地方.例如一个集合能够被分拆成的部分的数目不大于该集合的所有元的数目这个命题的正确性不能得到证明,除非通过把讨论的每一部分与它的一个元素联系起来."[5](这里"数目"是指"基数").如策梅洛所知,这个结果康托尔已经在19世纪80年代使用过.实际上在甚至更早的时候选择公理就被使用了,尽管没有被陈述.严格来说,对那些选择规则已具体确定的选择来说,这个公理不是必要的.但是早在1871年的时候,康托尔和海涅在他们宣传魏尔斯特拉斯未发表的研究的工作中,在不能具体给定一个选择规则的情况下来证明一个在点p序列连续的实函数(见18.2.2节)在点p也连续时已经隐含地使用选择公理.(在每一个特定的区间,一系列数必须被选定.)戴德金在选择关于一个理想的某些等价类的代表元的时候也隐含地用了这个公理.尽管这个公理已用了30年这个事实,策梅洛发表它后依然很快就引起了暴风雨般的争议.

争论的实质是无限多地任意选择的运用在数学上是否合理.这个问题很快就成为另一个更大的问题的一部分,即什么方法在数学上是完全许可的,以及是否所有的方法都必须是构造性的.然后当然地就有这些问题产生:什么组成这些构造?一个数学对象存在意味着什么?等等.数学家们以前几乎没有争论过这样的观点,但是看起来平淡的选择原理的运用现在却导致了一个结果的证明:良序定理,而对该定理许多数学家都是持怀疑态度的.因此数学界关于策梅洛的结果的有效性在观点上产生了很大的分歧.一些人完全接受了它们并继续在他们自己的研究中明确地使用选择公理,而另一些则否认这个公理及策梅洛的证明的有效性.一些数学家,罗素是其中的杰出成员,则采取

中间态度，谨慎地试着去判断哪些在集合论上是已经被接受的结果的证明依赖于选择公理，这样就可明确地知道拒绝这个公理将意味着什么.不幸的是在那时甚至连罗素也不能证明一个特定的结果其证明是否需要用到选择公理.

人物小传	恩斯特·策梅洛(1871—1953)(Ernst Zermelo)
	策梅洛作为一个教授的儿子，在柏林长大.他先后在柏林，哈雷和弗赖堡等地的大学里学习，最后以在变分法方面的一篇学位论文在柏林大学获博士学位.他在哥廷根当了几年教师，但在 1910 年移居到了苏黎世，恶劣的健康状况迫使他在 1916 年辞职，但出于对他在集合论上的重要工作的赞赏，希尔伯特帮助他获得了基金，这样他能够移居到黑森州进行一段长时期的修养.1926 年他返回到弗赖堡教学，但在 1935 年他因不能接受在大学里的新的政策而辞职.二次大战后他返回了弗赖堡，在那里，他完成了他的事业.

18.1.3　集合论的公理化

在这场数学界的争论中的一个问题是，没有一个被接受的集合论的公理集，使得人们可以通过其来判断什么方法是可接受的.那时发表了很多原理，特别是在策梅洛的证明后，一些数学家接受它们而其他人则否认它们.因此，策梅洛觉得为了巩固他的证明以及阐明证明中用到的术语，集合论应该是公理化的，这样他的证明能够被嵌进公理化当中.这个公理化不仅要包括选择公理，而且要设计一个公理来清除因康托尔对集合过宽的定义而产生的一些悖论.策梅洛的公理化方法受希尔伯特的几何公理化思想的影响，首先从一些不确定的对象的汇集着手，由公理给出它们之间的关系.换句话说，策梅洛首先从一些对象的范围 $\mathscr{B}$ 开始，然后给出任何两个对象的从属关系 $\in$.一个对象称为一个集合，是指它包含另一个对象(除去由公理 2 特别指明的).$A \subseteq B$，是指对任何 $a \in A$，都有 $a \in B$.策梅洛的 7 个公理如下(其名称是他给出的)：

1.(外延公理)对集合 S 和 T，如果 $S \subseteq T$，且 $T \subseteq S$，那么 $S = T$.

2.(初等集合公理)存在一个没有元素的集合，叫做空集.对 $\mathscr{B}$ 中任何对象 a 和 b，存在集合 $\{a\}$，$\{a, b\}$.

3.(分离公理)假如对集合 S，命题函数 $P(x)$ 是确定的(参见下面)，那么有集合 T，它恰好只包含了那些 $x \in S$ 使得 $P(x)$ 为真.

4.(幂集合公理)若 S 是一个集合，则 S 的幂集 $\mathscr{P}(S)$ 是一个集合.(S 的幂集即 S 的所有子集构成的集合.)

5.(并集合公理)若 S 是一个集合，则 S 的并是一个集合.(S 的并是指 S 中的元的所有元构成的集合.)

6.(选择公理)若 S 是非空集合的一个不交集，则有 S 的并的一个子集 T，它与 S 中的每一个成元恰有一个公共元.

7.(无穷公理)存在一个包含空集的集合 Z，使得对每一个对象 a，若 $a \in Z$，则 $\{a\} \in Z$.[6]

策梅洛从不谈及为什么他恰好选择如他所选的那些特定的公理，但人们可以推测他们中的大部分原因.第一个公理仅仅是断定一个集合由它的成员所确定，而第二个可能是起因于策梅洛对作为一个合法的集合的空集的需要以及区分一个元与仅包含这个元的集合.类似的，幂集合公理和并集合公理被设计用来阐明由其它集合所构造某些类型的集合的存在性，那些集合在许多证明中被使用.分离性公理是策梅洛用来纠正康托尔由任何性质定义集合的定义，从而消除罗素悖论，也就

是说,由这个公理,必须首先有一个描述性质的函数所应用于的给定的集合 S,其次有一个明确命题函数,这个函数由这样的方式来确定:由 $\mathscr{B}$ 上的从属关系和逻辑法则总可确定对 S 中的任何特定元 x 来说 $P(x)$ 是否成立.最后策梅洛设计了无穷性公理来阐明戴德金关于无限集合的存在性的证明.那个证明不为许多数学家所承认,部分原因是它看起来是一个心理学的证明而不是数学的证明.因而策梅洛提出了他自己的公理,断定一个无限集是可以构造的.

对策梅洛的公理化的反应是混合的.首先,策梅洛因为没有证明他的公理是相容的而被指责,毕竟,希尔伯特通过依赖实数的相容性对他的几何公理做到了一点.策梅洛承认他不能证明相容性,但是觉得它最终能够做到.无论如何,他相信从康托尔的所有集合理论都能由其推导出的这个意义上来说,他的体系是完备的.其次,策梅洛因他包含进的特定的公理和他排除的那些而受指责.对一个公理化体系的正确基础,当然没有一个一致的意见,并且由于策梅洛不能表明他的体系没有缺陷,就很难使数学界的大多数人相信这些公理将达到所需的目标.

为了获得一致认同,策梅洛体系不得不发生两处变化.首先,这些公理本身需要稍微的修改.在几位数学家的建议下,策梅洛自己在 1930 年引入了一个新的体系.现在叫做策梅洛 - 弗伦克尔集合论(因 Abraham Fraenkel(1891—1965) 而命名).相比于策梅洛最初体系的主要的变化是引入了一个新的公理:替换公理.设计此公理之目的是为了确保策梅洛理论中集合 $\{\mathbf{N},\mathscr{P}(\mathbf{N}),\mathscr{P}(\mathscr{P}(\mathbf{N})),\cdots\}$ 的存在,这里 $\mathbf{N}$ 是自然数集.弗伦克尔对这个公理的最初表述是这样的:"若 M 是一个集合并且 M' 是由于范围[$\mathscr{B}$] 中的一些对象替换 M 中的每一个成员而得到,则 M' 也是一个集合."[7] 作为第二个改变,一个"明确"的命题函数的性质必须被阐明,因为这对分离公理是本质的.后来发现这种阐释比起与集合论的关系来说与逻辑论的关系更加有关联.最后这种观点逐渐被接受,即公理化的集合论需要被作为逻辑领域的一部分.由于种种我们不能在这儿谈论的原因,甚至在今天还有不接受一个或几个策梅洛的公理的一些数学家学派.但是公平地讲,在以这些公理为基础在数学中取得的成功已经使大多数工作着的数学家相信这些公理为集合论形成了一个可工作的基础.[8]

选择公理本身,尽管可能是策梅洛公理中最有争议的,但它却不仅仅在分析中,而且在代数中有大量的应用.例如利用它可证明每一个向量空间都有一个基,每一个交换环的真理想都能扩展为一个极大理想.它在新兴的拓扑学中也被再三使用.那里作为许多结果之一,它提供了任何一族紧空间的积是紧的这一证明的基础.在数理逻辑的研究中它同样被证明是本质的.

源于选择公理的最重要的数学工具之一是极大原理,通常称为佐恩引理,它最后被证明与选择公理等价.尽管对极大原理有许多先行者,我们以佐恩(Max Zorn(1906—1993)) 在 1935 年给出的形式来叙述它:若 $\mathscr{A}$ 是一个集合族,包含的每一个链 $\mathscr{B}$ 也包含了它的并,则存在 $\mathscr{A}$ 中的一个集合 A^* 不是其他任何 $A\in\mathscr{A}$ 的真子集.(一个链 $\mathscr{B}$ 是指一个由集合构成的集,其中对 $\mathscr{B}$ 中的每两个集合 B_1,B_2,要么 $B_1\supseteq B_2$,或者 $B_2\supseteq B_1$.)实际上佐恩阐述这个公理的目的是取代代数的各种证明中的良序定理.因为后者尽管与他的公理等价,但它不是代数的,因为它在某种程度上是超越原理.无论如何佐恩的引理很快成为数学工具箱中的重要部分,并且被广泛应用于代数和拓扑中.事实上,1939 年它被出版在布尔巴基的巨著《分析基本结构的数学原理》(Elements de mathématique: Les structures fondamentales de l'analyse) 的第一卷中,并在整本书中被一直使用.



图 18.2 波兰邮票上的巴拿赫.

尽管选择公理被证明是有用的,它的一些结果还是不稳固的,而且完全出乎意料.在这些结果中,最令人惊奇的是巴拿赫 - 塔尔斯基悖论.这个悖论是由巴拿赫(图 18.2)(1892—1945) 和塔尔斯基(Alfred

Tarski)(1901—1983)于 1924 年首先发现的.他们用这个公理证明了任意两个不同半径的球在有限分解下是等价的.可以拿一个半径为一英寸的球 A 和有地球那么大的球 B,并把每一个都分别分成相同数目的块 $A_1,A_2,\cdots,A_m$ 和 $B_1,B_2,\cdots,B_m$,对每一个 i 有 A_i 等同于 B_i.尽管有像这样可用选择公理证明的结果,阐明这个公理关于集合论中的其它公理的确切地位是很有趣的.这个公理不能导出矛盾这一点当然不是明显的.

策梅洛意识到对他的公理的相容性的证明将会是极端困难的.尽管在 20 世纪的整个 20 年代,他和其他人都在考虑这个问题,直到 1931 年一个把大部分生活花在了普林斯顿高等研究院的奥地利数学家哥德尔(Kurt Gödel)(1906—1978)本质地证明了不可能存在这样的证明.实际上他证明了在任何一个包含关于自然数的公理的体系——比如,戴德金的公理能在策梅洛 - 弗伦克尔集合论中被证明——在这个体系中证明公理的相容性是不可能的.哥德尔还证明这个体系是不完备的,即在这个体系中存在可表述的命题,它们及它们的反命题均不可证明.在给出这些结果后,处理选择公理的惟一希望是证明它是相对相容的,即把它加到公理集中后不会导致任何在没有它时也不会出现的矛盾.哥德尔在 1935 年秋天的时候给出了这样的证明.在接下来的三年时间里,他还成功地表明在策梅洛 - 弗伦克尔集合论中连续统假设是相对相容的.

人物小传

尼古拉斯·布尔巴基(Nicholas Bourbaki)

布尔巴基是一群以法国人为主的数学家的共同名字.他们自从 19 世纪 30 年代中期以来一直致力于他们认为是分析的基础结构的东西的写作.他们的工作主要在以下六个方面:集合论、代数、一般拓扑学、实变函数、拓扑向量空间和积分.同时也发表了其他的领域的东西,包括交换代数,微分流形和李群.尽管在近年来布尔巴基的产出看起来似乎较其在 20 世纪五六十年代的节奏已经显著放慢了.

布尔巴基是由 H.嘉当,A.韦伊和其他正在参加庞加莱研究所里朱利亚(Gaston Julia)领导的数学讨论班的人所创建的.它最初的目的是重写经典法文著作《分析教程》(Cours d'analyse)的一个现代版本,这本教材从 19 世纪早期以来(以各种版本形式)使用于法国的大学中.从范德瓦尔登(van der Waerden)的《近世代数》得到启示,布尔巴基计划把一个给定的领域的重要定理建立在合理的公理集合上.他想陈述那些无论是古代还是近代的,只要曾导致重大应用和已被证明导出了重要概念的定理.

布尔巴基一般花十到十二年的时间来写一章或几章.一个成员被分派写初稿.大约一年后,初稿被拿到布尔巴基会议上接受详尽而毫不留情的审定.一旦对这个初稿有争议,另一个人就会被选择来修改它,然后下一年修改稿被重新讨论.但是最后布尔巴基在内容上达成一致意见,然后书被出版.

布尔巴基的著作以其绝对的严格和对广泛的一般概念的专注而著称.他们通常的方式是由这些一般思想开始,然后尽可能的由他们衍生,直到最后考虑那些经常历史性地促进一般理论的发现的特定例子.布尔巴基的影响及于美国的研究生院,特别是在 20 世纪的五六十年代,甚至对较低层次的数学教学也有影响.但是近年来甚至连法国的中学教育都抹去了布尔巴基的影响.

决定选择公理与策梅洛 - 弗伦克尔集合论的关系的最后结果由科恩(Paul Cohen)(生于 1934 年)在 1963 年完成.用全新方法,科恩证明了选择公理和连续统假设都是独立于策梅洛 - 弗伦克尔集合论(不含选择公理)的.换句话说,不可能在集合论中证明或否定这两个公理中的任何一个.进一步,人们可以自由地采用两者中的任何一个的反命题而不用担心会在集合论中引入新的矛盾.由

于这些及近年来的其他更多的结论,集合论中似乎有如在几何学中已经表明的那种情况,即依赖于各自的选择公理,会有不是一个而是许多个不同可能的版本.这对数学的进步是好还是坏是应该由下个世纪来判定的事情.

18.2 拓扑学

拓扑,作为几何中考虑连续并有连续逆的变换下图形的不变量的性质的部分,从 19 世纪的各种根基上在 20 世纪早期成长为了羽翼丰满的数学分支.这里我们只考虑这个学科的两个分支:点集拓扑学和组合拓扑学.点集拓扑学考虑某种抽象"空间"的点集的性质,组合拓扑考虑的是几何对象如何从有明确规定的"建筑块"来堆砌而成.

点集拓扑成长于康托尔对实数集的研究.就像这样,它的主要目的是提供一种合适的方法来推广像波尔查诺 - 魏尔斯特拉斯性质和海涅 - 波莱尔性质这样的实数的性质.这些都是和重要的现代的紧性概念紧密相关的.紧性概念是对闭区间上连续函数的最大值存在性定理推广的概念.波尔查诺 - 魏尔斯特拉斯性质是这样的性质:每一个有界的无限实数集至少包含一个**聚点**,即一个在包含它的每一个开区间都有集合中的另一个点的点,换句话说,这个性质断言了实数"完备性"的存在.我们在 16.1.3 节讨论了波尔查诺对这个结果的证明.

海涅 - 波莱尔性质,是 1894 年波莱尔(Emile Borel)(1871—1956)给出的,详细说明了若一个无限区间集 $\mathscr{A}$ 覆盖了一个有限实数闭区间 B,即 B 中每一个数都至少在 $\mathscr{A}$ 的一个元素的内部,则存在 $\mathscr{A}$ 的一个有限子集也有同样的性质.(尽管海涅在 19 世纪 70 年代已暗暗地使用了这个结果,但是对可数集合 $\mathscr{A}$, 正是波莱尔首先明确提出并证明了这个定理.1904 年勒贝格(Henri Lebesgue)(1875—1941)把这个结果推广到了任意的无穷集 $\mathscr{A}$ 上.)波莱尔的证明是用反证法.设 $\mathscr{A}=\{A_1,A_2,\cdots,A_m,\cdots\}$,如果结论不对,那么对每个 n,存在点 $b_n\in B$,使得对任何 $i\leqslant n$,b_n 都不在 A_i 中.假如把区间 B 分成两部分,则其中任何部分都有同样的性质.一直做下去,则得到一列递降的闭区间套 B_i,每一个都与 B 本身有相同的性质.但是 $\bigcap B_i$ 只包含一个点 P.由假设,P 是某个 A_k 的内点,所以 A_k 必包含一个区间 B_i,矛盾.

波莱尔的证明的关键是通常人们所指的区间套性质,即一个闭区间套的交包含一个点,正是这个结果后来被抽象成了最早的紧性定义,也是这个结果出现在把集合论作为一个整体的第一个系统讲述的文章中,那是 1906 年 W. 杨(William Young)(1863—1942)和 G.C. 杨(Grace Chisholm Young)(1869—1944)所著的《点集论》.

18.2.1 杨夫妇和《点集论》

杨夫妇的这本教科书研究的是实直线或实平面上的点集的性质,它给出了许多基本概念的明确定义,而这些概念后来又被更进一步推广.例如,属于一个两端均为"非闭"的区间的一个点 x 被称为相应的闭区间的一个**内点**.一个点 L 被称作一个给定的实数集合的**极限点**,是指在每一个包含 L 作为内点的区间内都存在属于集合的异于 L 的点.一个包含其所有极限点的集合被称为**闭的**,相反地则称为**非闭的**或**开的**.(注意这不是今天所用的"开"的定义)后来杨夫妇用这些定义重新叙述了波尔查诺 - 魏尔斯特拉斯定理和海涅 - 波莱尔定理并给出了证明.

接下来杨夫妇把直线上的"区间"推广到了平面中的"区域",他们的做法是把区域看成是由包含在其中的一系列三角形所生成.从而通过用区域代替区间把极限点的概念进行了推广.他们进一

步注意到,与区间的性质类似,一个区域把平面分成了三个分离的集合:内点(那些至少在一个三角形内的点),边界点(那些不是内点,但是内点的极限点)和外点(既不是内点也不是边界点的点).然后他们很容易地给出并证明了波尔查诺 - 魏尔斯特拉斯和海涅 - 波莱尔定理在平面上的推广.

现代拓扑的另一个基本概念,连通性,源于康托尔的一些想法.作为他处理"连续统",即整个实数集的工作的一部分,康托尔试图给出这个集合的特性.作为这个尝试的一部分,他发现确定此集合的"一片"是什么的含义是必要的.由于他按照集合中的点之间的最小距离为 0 这一点看到了这种思想,他按照距离来定义连通性:一个集合 T 是**连通的**,如果对 T 中任意两点 p,q 和任意一个正数 ε,均存在 T 中的有限个点 $t_1,t_2,\cdots,t_n$,使得距离 $pt_1,t_1t_2,\cdots,t_nq$ 均小于 ε.杨夫妇,及其他一些人,意识到给出一个不用距离,而是用纯粹的集合论术语的定义会更好,因此它们转述这个概念如下:"一个点集,以任何方式在集合的每一个点和每一个极限点周围形成一个区域并以这些点为内点,使得这些区域生成一个单一区域,若这个集合包含多于一个的点,则称其为**连通集**."然后用这个定义,他们证明了一个集合是连通的当且仅当它不能被分成两个无公共点的闭分支.

> **人物小传**
>
> 格丽丝·奇斯荷姆·杨(1868—1944)(Grace Chisholm Young)
>
> 格丽丝生于伦敦附近的哈斯勒米尔(Haslemere),在家中接受教育,然后进入了剑桥的格雷顿(Griton)学院,这所学院是英国第一个妇女能够接受大学教育的机构.1892 年在剑桥荣誉学位考试中获得优秀成绩后,她决定去哥廷根继续她的学业,因为在英格兰没有进行高级研究的可能性.F.克莱因愿意接收女学生,但是仅在他通过面谈确信她们会成功以后.(在系里有其他在任何情况下都拒绝接纳女学生的成员).格丽丝在 1895 年获得博士学位成为第一位通过正常步骤在数学上获得德国博士学位的女性.1896 年她与一个曾是她在格雷顿时的导师的英国数学家威廉·杨(William Young)结婚.
>
> 杨夫妇以伙伴关系度过了接下来的在数学和孩子(6 个)两方面都有丰硕成果的 44 年.尽管在接下来的超过 200 篇的数学论文和著作中大部分是署以 W.杨的名字,但格丽丝在他们的成果中扮演了一个重要角色.如威廉在 1914 年的一篇文章里所提到的,他和他的妻子讨论了工作的主要思想,并且格丽丝作了详尽的证明并把论文整理成可发表的格式.他们的女儿写道,他的父亲只有得到一个有共鸣的听众刺激才能够产生思想.格丽丝不仅提供了这个听众,而且她也有创造性和精力来完成正进行的各种他的丈夫提出来的工作.第二次世界大战爆发后威廉在瑞士的家中去世,而把格丽丝留在了英格兰.她去世于 1944 年,刚好在她即将从格雷顿学院获得荣誉学位之前.他们的两个儿子和一个孙女也成为了数学家.

18.2.2 弗雷歇和函数空间

就在杨夫妇论文发表的同一年,弗雷歇(Maurice Fréchet)(1878—1973)就着手将平面上点的一些结果推广到更一般的情形.在他的关于泛函理论的论文中就研究了什么时候两个函数相互"逼近"的问题,据此来"系统地建立泛函微积分理论的基本原理,然后将此原理应用于一些具体的例子."这里的泛函是指作用在函数集合上的函数.因为有了上述的工作,"人们常常受益,这体现在人们可以更清楚地看到在证明中哪些是本质的,体现在它的简洁明了和可以从只依赖于所考虑对象的特殊性质的证明方式中解脱出来."[10] 换句话说,弗雷歇决定用任意的集合重新考虑实直线上的拓扑的基本术语,然后把这些术语应用到他特别有兴趣的集合中.他发现可以在一般的情形下把拓扑学中的许多结论统一起来证明,然后将结论应用于具体的例子中,而不是将同样一个结论反反

复复地来证明.他尤其想解决的问题是关于泛函收敛于极限的问题,以及极限泛函与原来的泛函在什么条件下有相同的性质.

弗雷歇开始不是定义“极限”,而是用公理来描述极限.也就是说,一个集合 E 属于极限被定义的类 $\mathscr{L}$,是指对给定的 E 中任何无穷子列 $\{a_i\}$,都可以决定是否惟一地存在一个有一定性质的元素 a,这个元素是 $\{a_i\}$ 的极限.极限元素 a 将符合下列条件:如果每一个 $a_i = a$,那么极限是 a,且如果 a 是 $\{a_i\}$ 的极限,那么它也是 $\{a_i\}$ 的任何子序列的极限.对于极限这个抽象的概念,弗雷歇陈述了大量的定义,其中一部分也被康托尔考虑过:E 的**导出集** E' 是其所有极限元素的集合.如果 $E' \subseteq E$,那么说集合 E 是**闭的**;如果 $E' = E$,则说是**完备的**.若 a 不是 E 外任何序列的极限,则称点 a 是 E 的**内点**.

作为其他泛函研究的一部分,弗雷歇想对海涅 - 魏尔斯特拉斯关于闭区间上连续函数的最大和最小值结果进行推广,其目的在于把它应用于整个函数空间.为完成这个推广,1904 年他把海涅 - 波莱尔定理的证明中的中心思想在 1904 年的一个简短注释中当作了一个定义:“我们称一个集合 E 为**紧的**,是指对 E 中的每一个无穷子集序列 $E_1, E_2, \cdots, E_n, \cdots$,如果它们是闭的(并每个集至少含有一个点),且每个包含在前一个中,则至少存在一个公共元.”[11] 接着弗雷歇证明了集合 E 为紧集的一个充要条件是每一个无穷子集 F 至少有一个 E 的极限元,这个元是 F 中某个不同点的序列的极限.通过他的定义弗雷歇注意到紧集合具有与空间中有界闭集类似的性质.弗雷歇通过使用连续性的定义推广了魏尔斯特拉斯的结果,即得到了今天广为人知的序列连续性概念.一个函数 f 在闭集合 E 中的一点 a 处是**连续的**,是指对 E 中任何收敛于 a 的序列 $\{a_n\}$,都有 $\lim_{n\to\infty} f(a_n) = f(a)$.有趣的是,弗雷歇在他的 1906 年的论文中,把前面所提到的充要条件而不是相交性质看成为“紧”的定义,并且把任何有限集合也认为是紧的.

因已证明距离的概念对推进许多常用的概念没有必要,弗雷歇继续在更一般情形重新引进这个概念.他考虑类 $\mathscr{L}$ 的子类 $\mathscr{E}$,它由能够被定义度量的元素 E 构成.一个**度量**是指一个实值函数 (a, b),对任何三个元素满足:(1) $(a, b) = (b, a) \geqslant 0$;(2) $(a, b) = 0$ 当且仅当 $a = b$;(3) $(a, b) \leqslant (a, c) + (c, b)$.通过使用这种度量,弗雷歇定义了柯西序列的概念.一个序列 $\{a_n\}$ 称为**柯西序列**,是指对任何 $\varepsilon > 0$,都存在数 m,使得对任何 p,都有:$(a_m, a_{m+p}) < \varepsilon$.极限元和与它们联系的概念接着也根据度量和柯西序列得到了定义.特别地,弗雷歇考虑了一个**正规**集合的子类,即由完全的,可分的(包含一个可数的稠密子集)集合组成,其中每个柯西列都有一个极限(按现代的术语,称为完备).对这种类型的集合他可以证明海涅 - 波莱尔定理的一个推广:一个正规集合 E 是紧的当且仅当对于集合 $\{I\}$ 的每一个汇集 $\mathscr{G}$,如果 E 的每一个元至少是汇集 $\mathscr{G}$ 的一个成员的内部,则存在一个 $\mathscr{G}$ 的有限子集有相同的紧的性质.弗雷歇提供了大量具有度量的空间的例子用以说明他的定理的正确性,其中也包括许多函数集合的例子.一个例子就是给定闭区间上的实值连续函数的全体组成的集合,其度量是下面的极大范数:

$$(f, g) = \max_{x \in I} |f(x) - g(x)|,$$

在此极大范数下,弗雷歇证明了这个集合是正规的.第二个例子是所有的实数序列 $x = \{x_1, x_2, \cdots\}$ 组成的集合,其度量如下:

$$(x, y) = \sum_{p=1}^{\infty} \frac{1}{p!} \frac{|x_p - y_p|}{1 + |x_p - y_p|}.$$

弗雷歇注意到这种度量比更为标准的度量 $(x, y) = \max_p |x_p - y_p|$ 在某种意义上说更好.前者总给出一个有限的距离,而在后者却不这样.弗雷歇还证明了这个集合是正规的,而且对于一个具有合适度量的实值函数集合来说也是正规的.

18.2.3　豪斯多夫和拓扑空间

豪斯多夫(Felix Hausdorff,1868—1942)从实数集合的标准性质出发对拓扑空间的概念给出了完整的公理化描述.在他的 1914 年的著作《集合论基础》中,他描述了这个公理.虽然他的公理化和定义从 1914 年以来都被修改过且还引进了许多附属定义和概念.但现在普遍被教授的点集拓扑的基础在这本著作中都能找到.

豪斯多夫注意到三个基本概念:距离、邻域和极限.通过这三个概念能建立拓扑空间理论.弗雷歇曾有效地利用了第一和第三个.豪斯多夫注意到从距离概念出发可得出另两个概念,而若从邻域开始,则只能定义极限.然而一般来说,不能颠倒这些顺序.依豪斯多夫看,虽然一个人的研究方法仅是兴趣问题,但他仍打算从邻域概念出发去定义**拓扑空间**.这样的空间,今天称之为**豪斯多夫空间**.它是这样一个集合 E,其中每一个元素 x 都对应着 E 的子集族 $\{U_x\}$,这些子集族称为邻域,满足下面的公理:

1. 每一点 x 至少属于一个邻域 U_x,且每个邻域 U_x 都含有点 x.
2. 若 U_x, V_x 是包含点 x 的两个邻域,那么它们的交集也包含点 x 的一个邻域.
3. 若 $y \in U_x$,那么存在一个 U_y, $U_y \subseteq U_x$.
4. 对于两个不同点 x, y,存在两个邻域 U_x, U_y,它们的交是空集.

在证明对每个实数 ρ,由 $U_x = \{y \mid (y, x) < \rho\}$ 定义的邻域所生成的度量空间满足所给定公理后,豪斯多夫发展了与弗雷歇所作的相同的基本理论.其主要变化是,区域或以现代术语来说,开集是核心概念.对豪斯多夫来说,**区域** A 是一个仅仅包含内点的 E 的子集.而内点是由包含于 A 内的点 x 的邻域 U_x 来定义.整个 E 集合和每个邻域 U_x 由公理 1 和 3 都是开区域.豪斯多夫还证明了任意多个开集的并和有限个开集的交仍是开集.相反,一个**闭集**是包含它的所有的聚点的集合.豪斯多夫接着还证明了闭集恰恰是开集的补集.

利用弗雷歇对紧集的极限定义,豪斯多夫证明了区间套、闭集、紧集的性质.接下来又给出了海涅 - 波莱尔定理的推广,这仅对最初的证明作了细小的修改.为满足在任何拓扑空间都实用,他还对极限点和收敛性给出了很一般的定义:若 x 的每一邻域 U_x 包含除有限多个点外的 A 的所有点,则点 x 是无穷集 A 的**极限**.进一步说,因为任何 A 集要么有惟一的极限,要么就根本没有.在有极限的情况下,记为 $x = \lim A$,或者 A **收敛**于 x.[13] 其它的定义中还有连通性的定义:"一个非空集 A 是**连通**的,是指它不能分成两个非空的不相连的部分,这两部分对 A 来讲都是闭的.[即每个集都是所在空间 E 的一个闭集同 A 的交集.[14]]"豪斯多夫注意到,因为闭集是开集的补,所以他可能已经在定义中指出了连通分支应该是个区域.

因为邻域的思想是豪斯多夫拓扑空间发展的出发点,他利用了邻域概念定义连续性是毫不令人感到意外的.那就是说,他对实变量函数连续性采用了标准的 $\varepsilon - \delta$ 定义.并指出这个定义使用了直线上的邻域关系,紧接着再转变成拓扑空间的一般定义:"函数 $y = f(x)$ 在 a 点是**连续**的,仅当对点 $b = f(x)$ 的每个邻域 V_b,在点 a 存在一个邻域 U_a,其象在 V_b 中,即有 $f(U_a) \subseteq V_b$."[15] 豪斯多夫接着引出一个与之等价的定义:函数 $f: A \to B$ 在点 a 是连续的,当且仅当,对 B 的每一个以 $b = f(a)$ 作为一个内点的子集 Q,其原象 $f^{-1}(Q)$ 也包含作为内点的 a.使用这个定义,很容易证明,一个连续函数保持连通性和紧性.豪斯多夫自然地观察到了这些事实的第一点暗含着中值定理,而第二点则暗含了一个闭区间上的连续函数最大和最小值的存在性.换句话说,豪斯多夫证明了拓扑空间对实变量函数的那些经典结论是自然而然的.

18.2.4 组合拓扑

组合拓扑学是从空间曲面的多连通性思想的研究中产生出来的，这种思想由于黎曼对微积分的研究而逐步成型. 该思想得到了 19 世纪中期物理学家的进一步精确化，因为它在诸如流体动力学和电磁学领域中是非常重要的. 然而，正是贝蒂(Enrico Betti)(1823—1892) 在 1871 年通过使用无边界的类比于黎曼闭曲线的超曲面将多连通性思想推广到 n 维空间，此外，贝蒂还把此思想应用于 n 维空间里的微分形式的积分研究中. 两个曲面连通性的差别是一种判别曲面本质上“不同”的方法. 该方法还能判别是否存在从一个曲面到另一个曲面的连续的可逆函数. 然而，为获得区分曲面的方法，庞加莱发展了同调思想.

人物小传

亨利 · 庞加莱(1854—1912)(Henri Poincaré)

庞加莱(图 18.4) 像希尔伯特一样，是一名全能的数学家. 实际上他对数学的每一个领域都有贡献，包括物理学和理论天文学. 他出生于一个中产阶级偏上的家庭. 许多家庭成员为法国政府谋事. 庞加莱很小就对数学具有强烈的兴趣，并在法国中等学校所有学生数学竞赛中获得一等奖. 1873 年，他进入工科学校. 1879 年获得博士学位后便开始了其大学教学生涯. 首先在康(Caen) 大学，而后 1881 年在巴黎大学. 在他生命的后期，他转向普及工作，并写下了几本强调科学和数学的重要性的书，包括《科学和假说》、《科学的价值》以及《科学和方法》. 在《科学和方法》中，庞加莱描述了数学发现的心理学，强调了在数学创造中作为中心因素的下意识活动的重要性.

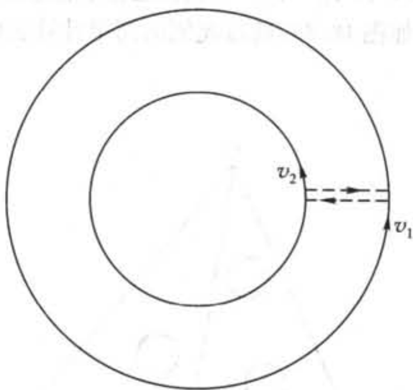


图 18.3 同调关系.



图 18.4 法国邮票上的庞加莱.

庞加莱在 1895 年的重要论文《位置分析》(Analysis Situs) 和四年以后的补充著作中，他作了如下的定义：在 n 维簇 V 的 p 维子簇 $v_1, v_2, \cdots, v_r$ 中若同调关系存在，记为

$$v_1 + v_2 + \cdots + v_r \sim 0,$$

仅当对某个整数 k，由所有 k 个 v_i 组成的集合构成了 $p+1$ 维子簇 W 的整个边界. (对庞加莱来说，簇(现在称做“流形”) 是对一维中的曲线或二维中的曲面的高维空间的推广. 一般来说，被想成，至少是局部被想成为，某些适当的函数方程组的零点的集合，或者是参数表示的某些函数的像集.) 庞加莱通过考虑定向引进了簇的“负性”，也就是说，$-v$ 与 v 是相同的簇，仅仅定向相反而已. 取一个同调关系的例子，设 v_1，v_2 分别是图 18.3 中带有特定方向的环的外边界和内边界，那么 v_1, v_2 共

同形成了该环的整个边界，且因为 $-v_2$ 与 v_2 有相反的方向，那么具有关系 $v_1 - v_2 \sim 0$. 庞加莱进一步发现，同调关系中的簇可以作加法、减法，以及与整数作数乘运算，因此，一组簇如果在整系数里无同调关系存在，则称该组簇是**线性无关**的.

为了弄清楚多重连通性的概念，庞加莱将簇 V 中的 p 维**贝蒂数** B_p 定义为比闭的 p 维线性无关子簇的最大数目还多 1，这里的闭簇是指一个无边界的簇. 这样以庞加莱来看，18.3 图中的环的贝蒂数为 2. 然而，双重环的贝蒂数却是 3. 另一方面，圆盘的一维贝蒂数为 1. 庞加莱将他的同调思想运用到各种维数的空间的积分研究当中，他还打算证明一个对偶定理，即对紧的连通的可定向的 n 维流形，关系 $B_p = B_{n-p}$ 成立，其中 $n-1 \geqslant p \geqslant 1$. 然而，庞加莱的证明甚至其定义，从现代读者的眼光看，有不足之处. 人们花了将近 20 年的时间去建立可以包含庞加莱基本思想的同调理论，且那些思想是严格合乎现代标准的.

现代同调理论在 20 世纪早期被几个数学家进一步发展. 对庞加莱思想的一个主要简化是把 p 维子流形不是看作方程组的解，而是把它看作从一简单 p 维子流形形成的，每一个这样的 p 维子流形是一个 p 维"三角形" 的连续映象. 确切的定义被亚历山大(J. W. Alexander) 在 1926 年所解决. 当时，他把 p - **单形**定义为一个"三角形" 的 p 维类比，把**复形**定义为有限个单形的并集，这些单形之间任何两个都没有共同的内点，而集合里每个单形的面也是集合的单形. 复形的一个初等 i - 链定义为形如 $\pm V_0V_1\cdots V_i$ 的式子，这里的 V 是一个 i - 单形的顶点. 这个式子中如果顶点移位，则符号要改变，这样就给每个链一个定向. 每一个初等 i - 链都是一个 p - 复形的 i - 维"面". 而每一个 i - 链都是初等 i - 链的整数系数的线性组合. 例如，以 V_0, V_1, V_2, V_3 为顶点的四面体是一个 3 - 单形，而它与它的四个面(每个是 2 - 单形)，它的四个边(每个是 1 - 单形)，以及它的四个顶点(每个是 0 - 单形) 共同形成一个复形. 面 $V_0V_1V_2$ 是 3 - 单形的一个初等 2 - 链. 亚历山大接下来把初等 i - 链 $K = V_0V_1\cdots V_i$ 的边缘定义成$(i-1)$ - 链 $K' = \sum(-1)^s V_0\cdots\hat{V}_S\cdots V_i$，且把此线性地推广到任意 i - 链上. 这样 $V_0V_1V_2$ 的边就是 $V_1V_2 - V_0V_2 + V_0V_1$(如图 18.5). 通过此例的简单计算表明边缘的边缘是 0，且还能证明这个结论是普遍正确的.

亚历山大把他的同调定义运用到**闭链**上，即边缘是 0 的链上. 那就是说，一个闭链 K 同调于 0，$K \sim 0$，仅当 K 是一个链 L 的边缘. 两个链 K 和 K^* 是同调的，$K \sim K^*$，仅当 $K - K^*$ 同调于 0. 一复形的第 p 个贝蒂数那么就是那些 p 次闭链的最大数目，使得这些 p 次链关于边缘是线性独立的，那就是说，任何线性组合不同调于 0.(注意这个数比庞加莱最初定义的数要小 1.)

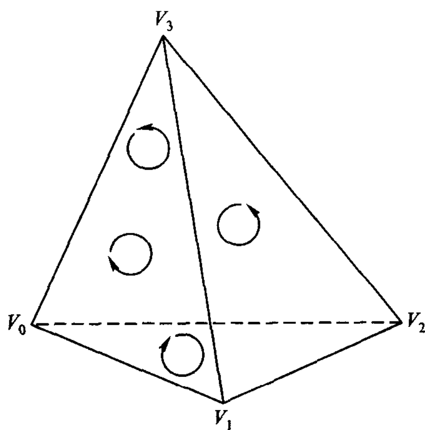


图 18.5　四面体的边缘.

随着在闭链上考虑带逆运算的可交换运算("加法")，对现代读者来讲，应该清楚在亚历山大定义中所隐藏的群的理论. 20 世纪 20 年代的数学家也看到了这点. 但是在讨论将群理论运用到拓扑学之前，我们首先需要的是去考虑在 20 世纪的初期一般的代数学的发展历程.

18.3 代数方面的新思想

人物小传	列昂纳德·欧仁·迪克森(1874—1954)(Leonard Eugene Dickson)
	迪克森出生于独立的依阿华州,是芝加哥大学数学博士学位的首位获得者.之后,他又到德国莱比锡和法国巴黎学习.然后在美国德克萨斯任教一年.最后于1900年回到了芝加哥,并一直在那里工作.迪克森是一位多产的富于创造力的数学家,他发表了几百篇论文和出版了大约18本著作.其中最重要的著作是三卷本的《数论的历史》.在这本著作中,他跟踪了数论领域中每一个重要概念的演变.迪克森从1911年到1916年担任"美国数学会会刊"的编委,并在接下来的两年任主编.1913年他当选为美国国家科学院院士.

希尔伯特发展了几何上的一组公理,并证明了它们的相容性,至少是归结为算术的相容性.自从穆尔(H.Moore,1862—1932)在1902年表明希尔伯特的公理不是相互独立的,即是说它们中的一条实际上能被另外的公理推出,在代数方面为发展出大量的独立公理已做出了巨大的努力.例如,在1903年迪克森(1874—1954)发展了一组新的关于域的公理.他认为那是对早在10年前韦伯的工作的一个改进,迪克森把具有两种合成规则的集合定义为域.这两种合成规则记为"+"和"×",满足下列的九条公理:

1. 若 a,b 属于此集,则 $a+b$ 也属于此集;
2. $a+b=b+a$;
3. $(a+b)+c=a+(b+c)$;
4. 对此集中任何两元 a,b,存在此集中元 x,使 $(a+x)+b=b$;
5. 若 a,b 属于此集,则 $a\times b$ 也属于此集;
6. $a\times b=b\times a$;
7. $(a\times b)\times c=a\times(b\times c)$;
8. 对此集中任何两元 a,b,使得至少存在此集中一个元 c,满足 $c\times a\neq a$,则在此集中存在一个元 x,使得 $(a\times x)\times b=b$;
9. $a\times(b+c)=(a\times b)+(a\times c)$.

自然,公理中的加法和乘法的封闭性、交换律、结合律还有分配律都是大家熟悉的.迪克森对加法和乘法的革新是以新公理4和8代替了单位元和逆元的规则.然后他继续进行证明其独立性,方法是对每条公理去创造一个具有不满足此公理但满足其余八个公理的具有两种合成规则的系统.例如,满足普通乘法规则和一个新的加法规则 $a+b=b$ 的所有正有理数的系统满足除第二条公理以外的所有公理.类似地,具普通乘法和给出的 $a+b=-a-b$ 的加法规则的所有有理数集合满足除第三条公理的所有其他公理.此时,公理4由取 $x=2b-a$ 所满足.

18.3.1 p 进位数

由于许多类似于迪克森的那些论文都涉及到处理像群、逻辑代数和线性结合代数(将在18.3.4中讨论)这些结构,代数的公理化方法受到了偏爱.一旦清楚了某个特别的构造是重要的,数学家们就会想方设法去发现一些独立的公理组而不诉诸于公理组的一个特别具体的表现形式.然而,为了得到新

的尚待证明的猜想,具体的例子总是需要的.固此,在20世纪初,亨泽尔(Kurt Hensel,1861—1941)就发现了一种新型的域,即 p 进位数域,它不同于韦伯所了解的那些 19 世纪的域的例子.

亨泽尔的出发点是注意到,若给出任一素数 p,一个正整数可惟一地写成

$$A = a_0 + a_1p + a_2p^2 + \cdots + a_rp^r,$$

其中每个 a_i 满足 $0 \leqslant a_i \leqslant p-1$.在此新表示中,两个数若模 p 的高次幂同余,则被想像成相对于 p 是"靠近"的.为更方便地发展他的理论,亨泽尔运用这个表示与分数的通常十进制表示的类比把 A 写成 $A = a_0.a_1a_2\cdots a_r$ 的形式,其中 p^k 的系数位于小数点后的第 k 位.于是,分别由数 3.2 和 3.2000000004 表示的两个数 $3+2\cdot5$ 和 $3+2\cdot5+4\cdot5^{10}$,如果被想成是十进位小数是接近的,也可以想成是相对于素数5是"接近"的,因为它们模5的直到10次的幂都为同余.更明确地,亨泽尔把数 $A_k = a_0.a_1\cdots a_k$,$k < r$ 称为 A 相对于 p 的近似值,这是因为模较 A_{k-1} 更高的 p 的幂时每一 A_k 与 A 同余.

要把这个新表示下的正整数转化成一个域,人们需要在它们上面进行通常的运算.按照有限(和无限)十进位数类比的方式作为引导,亨泽尔指出人们如何必须这样做.用"十进位"的类比,加法和乘法几乎是按标准的方式进行,但不同的是人们要从左到右进行运算,并且当和或者积大于或等于 p 时推进 p 的一个适当的倍数.举例来说,仍令 $p = 5$,便有下面的和,

$$\begin{array}{r} 2.3042134 \\ +\ 3.2413123 \\ \hline 0.10113031 \end{array}$$

像所预料到的那样,正是减法和除法运算迫使亨泽尔在这个正整数集中添加进了新的元素,这些元素以一种令人惊讶的方式表示.考虑减式

$$\begin{array}{r} 3.131312 \\ -\ 4.424322 \\ \hline \end{array}$$

从左端开始,如有必要则向右面借数,则得到答数的前六位为 4.10243. 换句话说,对 $j \leqslant 5$,有 $3.131312 - 4.424322 \equiv 4.10243 (\bmod 5^j)$. 但是,能够得到一个精确的答数而不仅仅是个近似值吗?毕竟,如果把这些表示能换回通常的十进位表示,则答数是一个负整数.为了引进这些"负"数,亨泽尔只简单地允许小数点右进的数位无限地增加下去.在这个例子中,可在减数和被减数的右端放上无限多个 0,从而可以看出答数为 4.102434444…,它的一层意义在于,如果在第 n 数位上截断此数时,所得的数与实际的差数模 5^n 同余.亨泽尔进一步指出运用小数点后无穷多的位数对于在 $B \not\equiv 0 (\bmod p)$ 的特殊情形下能进行除法 $\dfrac{A}{B}$ 也是必要的,而且在减法与这种情形的除法下,这种无限 p 进制的展开式是循环的.

在把有理数域的一个子域(即那些商数 A/B,其中 $B \not\equiv 0 (\bmod p)$)中的元重写成一个循环的 p 进位展开式后,亨泽尔引进了一个新的集合以作类比于普通十进位表示的推广;这个集合是所有如下面形式的幂级数:

$$\begin{aligned} A &= a_0 + a_1p + a_2p^2 + \cdots + a_np^n + \cdots \\ &= a_0.a_1a_2\cdots a_n\cdots, \end{aligned}$$

其中每个系数 a_i 满足 $0 \leqslant a_i < p$.像前面一样,因为 $A_k = a_0.a_1a_2\cdots a_k$,$k = 1,2,\cdots$ 模 p^{k-1} 同余于 A,故而有限级数 A_k 的每一项都近似于 A.现在由于有了无穷多个这种近似值,而且每个都较前一个更好,亨泽尔便转向了分析,写下了

$$A = \lim_{k\to\infty} A_k.$$

但是如此定义的幂级数的集合并不是一个域. 尽管人们可以在各种近似表示上运用早先给出的规则进行加、减、乘以及取极限的运算,但要去除一个任意的幂级数就必须把这些级数拓展到含有有限项负指数的级数.(一个极简单的情形是,$6 \div 5$ 可写为 $1 \cdot 5^{-1} + 1$.)因此亨泽尔加上了这样的项,并且容易证明所有形如

$$A = a_m p^m + a_{m+1} p^{m+1} + \cdots \quad (m \text{ 为任意整数})$$

的级数集合确是一个域,即今天被称为 p 进数域 $\mathbf{Q}_p$.

现在已定义了一个新的域的集合,即对应每个素数 p 有一个域;在这之后,亨泽尔便能把域的一般理论中的各种概念应用到这些特别的域上了. 因此,这种域的整数便是那些 p 的最小幂次为非负的级数,而单位元则是那些 p 的最小幂次为零幂的元.(确切地说单位元是那些乘法的逆元仍是整数的整数.)亨泽尔处理了系数在这种域中的多项式,并将通常的构造方法应用到它们,这包括把这些多项式的根添加到域中以形成扩域,从而可以应用伽罗瓦理论. 但是,因为 p 进数域有一个自然的拓扑结构,即用同余于一个给定的 p 的幂来定义一个点的邻域,所以人们既可使用分析的概念也可使用代数的概念. 例如,p 进数的柯西序列的观念和极限的观念都可以得到定义,从而能够证明每个这样的柯西序列有一个极限.

18.3.2 域的分类

受到亨泽尔定义新的域的创造性工作的影响,斯泰尼茨(Ernst Steinitz,1871—1928)着手对整个域论进行全面的研究. 他的研究成果在 1910 年以专著《域的代数理论》问世. 斯泰尼茨的目标是"综述所有可能类型的域,并建立它们之间的关系"[16]. 其中最引人注目的是素域的概念和特征. 任何域 K 的**素域**就是其最小子域,也就是说,它是所有子域的交集. 若 ε 是该素域的乘法单位元,那么对于 ε 的所有整数倍 $m\varepsilon$ 的集合 I 有两种可能性. 首先,他们都是互不相同的,此时,I 与正整数集合同构,该素域就与有理数域同构,就称 K 具有**特征** 0. 其次,有一个最小的素数 p,使得 $p\varepsilon = 0$,此时 I 与模 p 的剩余类同构. 这样,I 本身是一个域,它是有 p 个元素的有限域,我们称 K 具有**特征** p.

斯泰尼茨通过讨论各种类型的扩张域而继续他的工作. 特别地,如果域 K 的扩张域 L 中的每一个元素都是系数在域 K 中取值的多项式方程的根,那么 L 称为**代数的**. 否则,就称为**超越**扩张域. K 上的一个代数扩域 L 说是 n **维有限**的,是指 L 中存在 n 个在 K 上线性独立的元素,而任何超过 n 个元素都是线性相关的. 若线性独立元素的个数不是有限的,则扩域称为是**无限**的. 有限的代数扩张域进一步被分成两部分:一部分适用于伽罗瓦理论;另一部分则不适合.(伽罗瓦理论从完整的意义上讲,仅适用于那些对不可约多项式有不同根的域.)对于伽罗瓦理论适用的域,包括所有具有 0 特征的域,称为是**可分**的;而其它的域则称为**不可分**的. 例如,无限的代数扩张域包括一个素域 K 的**代数闭包**,其中 K 上的每一个多项式都可以分解成线性因子. 斯泰尼茨给出了对具有 p 特征的素域,怎样去构造代数闭包的方法,但是也指出对有理数域,这个构造涉及到选择公理的使用. 那么毫不惊奇,对于任意域的代数闭包的存在性的一般证明也涉及到那个公理. 最后,斯泰尼茨讨论了超越扩张域. 由于他把一个域 K 的纯超越扩张域定义成由添加有限或无限多个未知量形成的域,并使用了良序定理,因此,斯泰尼茨证明了每一个 K 域的扩域能够由取一个纯超越扩张,然后再取一个代数扩张形成.

18.3.3 向量空间的公理化

斯泰尼茨在他的域的有限代数扩域的定义中使用了线性代数. 且他以域的扩张思想,继续去证

明了有限维向量空间的一些结果.特别地,他精确地证明了早些年前戴德金含蓄的假定:对于一个有限维的向量空间的生成集不可能有比空间维数更少的元素.他的论述和该结论的证明当然是依据由生成元所决定的坐标系得到的.因为假设域 K 上的空间 L 的维数是 n,这样就存在 n 个线性独立的生成元$\{\alpha_i\}$.另外,如果存在一组 $n-1$ 个生成元 $\beta_j, j=1,2,\cdots,n-1$,那么每一个 α_i 就能写成形式 $\alpha_i = c_{i1}\beta_1 + c_{i2}\beta_2 + \cdots + c_{in-1}\beta_{n-1}$,这样方程 $d_1\alpha_1 + d_2\alpha_2 + \cdots + d_n\alpha_n = 0$ 就有一个非平凡解$\{d_i\}$.因为这个方程与下面的方程组等价:

$$\begin{array}{ccccccc} c_{11}d_1 & + & c_{21}d_2 & + \cdots + & c_{n1}d_n & = & 0, \\ c_{12}d_1 & + & c_{22}d_2 & + \cdots + & c_{n2}d_n & = & 0, \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ c_{1n-1}d_1 & + & c_{2n-1}d_2 & + \cdots + & c_{nn-1}d_n & = & 0. \end{array}$$

且任何比这个方程组有更多未知数的齐次方程组都有非零解.这样 α_i 是相关的,与最初的假设矛盾.

虽然斯泰尼茨为有穷维向量空间理论总结出许多基本的结论,但它们都是依据域的代数扩张得出的,他没有打算去对一般的向量空间给出一个公理系统.像我们以前指出的那样,早在 20 年前的佩亚诺系统被人们忽视了.而正是外尔(Hermann Weyl)(1885—1955),在他 1918 年的著作《空间 - 时间 - 物质》中,试图给这一领域进行公理化处理,以期从基本原理出发来为相对论的发展打下一个基础.尽管没有迹象表明他熟悉佩亚诺的工作,但他们的公理化系统本质上是相同的.惟一的不同就是,他不像他的前辈,他坚持认为向量空间是有穷维的.就这样,他的最终公理这样陈述道:"存在 n 个线性独立的向量,但是任何 $n+1$ 个元素都是线性相关的."[17] 很不幸,外尔的著作比起佩亚诺来,影响甚至更小.就这样,关于向量空间的思想就需第三个人去发现,这一次依据的是分析学.

在 20 世纪 20 年代,有几个数学家对他们领域中涉及的代数和拓扑性质的概念的推广很感兴趣.他们意识到弗雷歇的度量空间思想同向量空间的思想是相似的.为把这两种思想统一起来,巴拿赫在他 1920 年的论文中,引进了现在被称着**巴拿赫空间**的概念.所谓的巴拿赫空间,就是具有范数(一种距离函数)的向量空间,在此范数下,所有的柯西序列都收敛.就像他在 1922 年发表的文章中写到:"目前工作的目标是去建立确定的理论,该理论要在不同的函数领域都要有效 … 然而,没有必要去证明对每一个特别的领域成立,因为那样是痛苦的.我已选择了一条不同的路线.那就是,我将在一般意义上考虑几组集中的元素.这些元素我先假定它们有特定的性质,然后我将从它们中得出某些定理.接下来再证明对于每一个特定的函数领域,假定都是正确的."[18] 就这样,巴拿赫就在实数域内得出了能刻画向量空间特性的包括 13 个公理的系统.且因为巴拿赫对函数空间感兴趣,他的向量空间就没限定在有穷维内.虽然巴拿赫的公理系统含有多余的公理,但他的论文有巨大的影响.当《线性算子理论》出版 10 年以后,此书中的公理仍被重复应用着,从此,向量空间的抽象观念成为了数学词汇的一部分.

18.3.4　环　论

斯泰尼茨和戴德金的向量空间也是域,这样,除加法运算外,它还具有合理的乘法运算.具有两种运算加法和乘法的数学结构 R 现在被称作是**环**,是指对于加法 R 是交换群,对乘法满足结合律,分配律也成立.在**线性结合代数**的名义下,对该类结构的详细研究由美国数学家皮尔斯(Benjamin Peirce)(1809—1880)在 1870年所开创.线性结合代数今天仅被称为**代数**.皮尔斯认为一个环同时也是域 F 上的一个有限维向量空间.(皮尔斯把系数域限定在实数域.)他工作中的主要目标是通过

考虑基本元素可能的乘法表来描述所有 1 维到 5 维的代数及部分 6 维代数. 他的任务没有完成, 然而他引进了两个重要的定义: 如果环 R 中的一个非零元素的幂 a^n 等于 0, 那么称此元是**幂零**的, 而如果 $a^2 = a$, 则称是**幂等**的. 皮尔斯接着证明了下面的定理:

定理 在每一个代数中, 至少有一个幂零元或者一个幂等元.

这个定理的证明并不困难. 因为代数是有限维的, 代数中的任何非零元素 A 一定满足形如

$$\sum_{i=1}^{n} a_i A^i = 0$$

的方程. 此方程可重写为 $BA + a_1 A = 0$ 或 $(B + a_1)A = 0$, 在此 B 是 A 的幂的线性组合. 对每一个 $k > 0$, 有 $(B + a_1)A^k = 0$, 因此 $(B + a_1)B = 0$ 或 $B^2 + a_1 B = 0$. 如果 $a_1 \neq 0$, 则从上面最后一个方程立即有

$$(-B/a_1)^2 = -B/a_1,$$

于是 $-B/a_1$ 是幂等的. 如果 $a_1 = 0$, $B^2 = 0$, 那么 B 是幂零的.

在 19 世纪最后 25 年里, 其他几位数学家研究了特殊的代数, 尤其在**单代数**方面, 这是没有非平凡的双边理想的代数. (代数或者任何环 R 的一个**双边理想**是其一个子集 I, 满足: 若 α, β 属于 I, 则 $\alpha + \beta$ 也属于 I, 对于 R 内的任何 γ, $\gamma\alpha$ 和 $\alpha\gamma$ 也属于 I. 这个理想推广了戴德金对代数整数环中理想的定义. 自然而然, 双边理想也可以看作子代数.)

E. 嘉当证明了每一个(有乘法单位元的) 复数上的单代数是一个矩阵代数, 即是说, 对某个 n, 它与复系数的 $n \times n$ 矩阵代数同构. 他的工作后来被韦德伯恩 (Joseph Henry Maclagan Wedderburn) (1882—1948) 在 1907 年的论文《超复数》中推广, 这篇论文对任意域上代数的结构作了详细的讨论, 其中一个结果就是韦德伯恩证明了任何单代数是矩阵代数, 它不必在域上, 而是在可除代数上. (可除代数是带有乘法单位元的代数, 它的每一个非零元素都有一个乘法逆元素.)

为进一步划分代数, 去划分可除代数是必要的. 弗罗贝尼乌斯已经证明在实数域上仅有 3 个可除代数, 实数、复数和四元数. 韦德伯恩自己在 1909 年证明了仅有的有限可除代数本身是有限域, 且这些域是已知的. 在 p - 进数域上的可除代数被哈塞 (Helmut Hasse) (1898—1979) 在 1931 年分类. 而任何代数数域上的可除代数在 1932 年被哈塞、布劳尔 (Richard Brauer) (1901—1977) 和诺特 (Emmy Noether) (1882—1935) 所划分. 但是这些划分涉及到代数数论上的一些专题, 如类域理论, 这里我们不作讨论.

人物小传	约瑟夫 · H. M. 韦德伯恩 (1882—1948) (Joseph H. M. Wedderbarn)
	韦德伯恩出生于苏格兰的福法尔, 此地位于爱丁堡的北部 50 英里. 他在芝加哥大学向迪克森学习一年后, 在爱丁堡大学获得了学位. 在取得博士学位后, 在 1901 年他被威尔逊 (Woodrow Wilson) 邀请到普林斯顿, 加入到他新组建的教师队伍中. 由于年轻的教授们能提供小班指导, 普林斯顿变得著名起来. 在第一次世界大战中, 韦德伯恩加入了英国军队, 在法国作战之后回到普林斯顿, 在那里, 他负责《数学年刊》, 直到 1928 年. 在 20 世纪 20 年代晚期, 他明显遭受着神经性损伤. 之后, 他过着越来越孤独的生活, 很早就从教授职位上退了下来.

18.3.5 诺特环

环论上的许多工作都是哥廷根的诺特和其他数学家在 20 年代完成的. 诺特对理想创立分解理论 (它类似于戴德金的素分解, 但是前者适合于一般的环), 比其前辈的代数域里的整数环更具一般

性. 这些一般的环现在被称为**诺特环**. 所谓的诺特环是具有单位元的交换环，满足升链条件，即对环中的理想链 $I_1, I_2, \cdots, I_k, \cdots$，若有 $I_k \subset I_{k+1}$，则这一关系在有限项后终止. 对这些环诺特得到了分解结果，这些结果比理想的惟一素分解要弱.

诺特通过一组公理刻画了满足素因子分解的整个戴德金理论的环所具有的特征：

1. 环 R 满足上升链条件.
2. 对每一个非零理想 A 的剩余类环 R/A 都满足降链条件.（降链条件与升链条件是一样的，只是用 $\subset$ 代替 $\supset$）.
3. R 有一个乘法单位元.
4. R 是一个整环，也就是说，它没有零因子.
5. R 在其商域是整闭的. 换句话说，域的每一个满足环上多项式方程的元素本身就在环中.（校注：此多项式的首项系数为单位元.）

接下来她证明了，如果 R 满足这五个公理，则在其整环的有限可分扩张域中的整闭包也满足这五个公理. 特别是，因为任何**主理想整环**（即每个理想均为主理想的整环）都满足这些公理，所以她的结论不仅表明在有限的代数数域中所有的整数环中的理想都有惟一的素分解，而且还表明在一个变量的代数函数域上的有限代数扩张中的整闭子环也具有此特性.

像这章开头描述的那样，诺特对其合作者有巨大的影响力. 特别在强调代数的结构方面而不是计算方面. 事实上，20 世纪的前 50 年内最重要的代数教程，由 B. L. 范德瓦尔登（1903—1996）写的《近世代数》的第二卷，主要归功于她，从其与早几年出版的教程相比较，就可以看出她开创的巨大变革（见补遗 18.2）.

人物小传

埃米 · 诺特（1882—1935）（Emmy Noether）

诺特得到了一个中上阶级德国犹太女孩的正规的教育，她完成了整个学业教育，学钢琴，上舞蹈课. 直到 1900 年，即在进一步学习法语和英语之后，她才通过巴伐利亚州的考试，有资格在中学任教. 但就在同时，她的兴趣从语言方面转向到了数学，接下来她花了三年时间在爱尔朗根大学旁听数学课程，她的父亲是这所大学的数学教授. 事实上，在第一学期，她是被允许旁听课程的两个女性之一. 1904 年，此大学正式招收女生时，她成了一名正式的学生，四年后，她获得了博士学位，其学位论文研究的是关于双二次型的不变量.

诺特在爱尔朗根大学又呆了几年，直到 1915 年，希尔伯特邀请她去哥廷根辅助他在广义相对论方面的研究. 虽然作为一名妇女她不允许去正式任教且不允许获得薪水，但希尔伯特仍安排她去教以他自己名义安排的课程. 事实上，他为她在大学理事会中辩论道："我真的不明白侯选人的性别就是一个被反对接受为无薪教师的论点. 毕竟，理事会不是浴室."[19] 直到一次大战后，德国这种状况得以改变，她才在大学里得到一个正式的位置. 但在 1922 年后，她才获得一份很少的薪水. 接下来在哥廷根的十年里，她的影响，不仅在德国，而且在苏联，都是巨大的. 她于 1928—1929 访问了莫斯科，1932 年，她是惟一一位被邀请到苏黎世的国际数学家大会上做特邀报告的女性.

她的处境像许多德国数学家一样，在 1933 年早期，由于纳粹党的上台而突然改变了. 作为一个犹太人，她失去了在哥廷根任教的机会，和她同事在一起被驱逐到国外. 她在靠近费城的布林马尔学院为自己找到了一个栖身地. 从 1933 年秋季开始，她就住在那里，此地离普林斯顿很近. 这为她定期去高等研究院参加学术活动提供了方便. 1935 年的 4 月，在一个貌似成功的肿瘤移植手术之后，她突然去世了.

<table>
<tr><td rowspan="3">补遗
18.2</td><td>数学中的女性</td></tr>
<tr><td>细心的读者应该已发现，在这本书中很少的女数学家被谈及到．当然，其原因是直到最近，从事数学研究的女性很少．或许在古代对数学有过贡献的女性其名字被历史遗忘了，但在有纪录的历史中，不管是在西方还是非西方文化中，一般地，妇女获得教育而在数学领域取得成功的条件不是那么好．就像本书中所涉及的女数学家，她们中的大部分都有一个亲密的家庭成员愿意教她们数学，或者至少鼓励她们学数学．如果没有这种条件，很明显妇女们很难跨入数学领域，即使她们真的愿意多学些数学知识，她们往往还是不能融入数学界．女士们被认为不适宜从事如此高智力的活动．</td></tr>
<tr><td>在过去的几十年里，这一状况正在发生变化．即使仍旧有巨大的障碍需要妇女们去克服，特别是对妇女的教育态度，但现在对那些想成为一个成功数学家的妇女们来说，已经是可能的了，即使没有家庭成员做榜样也行．事实上，在近几年，在美国授予的数学博士学位中大约 20% 授给了女性，在数学界女性担任重要位置的人正越来越多．在 80 年代美国数学会有过它的第一位女主席鲁宾逊(Julia Robinson)．美国数学协会近来也有几位女主席，还有一位女总干事．而妇女数学协会在过去的 25 年里积极地为妇女争取越来越多的机会．例如，为女研究生和博士生争取财政上的支持，赞助突出的女数学家在重要数学会议上做报告．在最近几次国际数学家大会上作报告的女性的人数越来越多，虽然这个数目仍旧很低．但无论怎样，明显地为妇女进入数学家这个职业所提供的机会显然有了进步．21 世纪末写下一本数学历史书时，将会有比这本书提及的更多的女数学家．</td></tr>
</table>

18.3.6 代数拓扑

1926 至 1927 年间诺特由于在哥廷根听了亚里克山大罗夫(Aleksandrov)(1896—1982) 关于拓扑学方面的系列演讲，她受到启发，产生了一个想法．这个想法开创了一个全新的研究领域．正如亚里克山大罗夫在纪念诺特的演讲中说的那样："在她参加我们的报告会中，当她熟悉了组合拓扑的系统的结构时，她就立即注意到直接研究给定多面体的代数复形和闭链的群以及所有零同调的闭链组成的链子群是值得的，而不是研究通常的贝蒂数和挠系数．她建议把贝蒂群定义成所有闭链组成的群与零同调的闭链组成的子群的商群"[20]．随着诺特的评论和后来韦托列斯(Leoplod Vietoris)(1891—?) 和霍普夫(Hopf)(1894—1971) 的相关文章的发表，代数拓扑开始变得热门起来．在 1927 年韦托列斯把一个复形 A 的**同调群** $H(A)$ 定义成模边缘的闭链群，这与诺特所提出的一样．几乎同时，霍普夫定义了几个其他的交换群．也就是说群 L^p，Z^p，R^p 和 $\bar{R}^p$ 它们分别是由 p - 复形，p - 闭链，p - 边缘(这些链是某一个链的边界)和 p - 边缘的因子(其倍数为边缘的链)产生的．接着对霍普夫来说，因子群 $B_p = Z^p/\bar{R}^p$ 是一个自由群(该群中没有一个元素的倍数是 0)，它的秩(基元素的个数)是该复形的第 p 个贝蒂数．

此新领域的进展是如此的快速，以至一年后，梅尔(Walther Mayer)(1887—1948) 为同调群的定义就发表了一个公理系统．也就是说，梅尔不再关心拓扑复形本身，而是关心定义在其上的代数运算．这样一个复环 Σ 是元素 $K^{(p)}$(复形)的一个汇集，其中每一个元附有一个维数 p，所有的 p 维元素形成一个自由的阿贝尔群 K^p．对每一个 p 都有一个同态 $R_p: K^p \to K^{p-1}$，使得 $R_{p-1}(R_p(K^p)) = 0$．(R_p 被叫做第 p 个边缘算子，通常人们使用不带下标的 R，于是上一个方程式可以写为 $R^2 = 0$．) 有了这些公理以后，梅尔把 p 次闭链群 C^p 定义成 K^p 的元素 K，使得 $R(K) = 0$，而把 p 的边缘群定义成 $R(K^{p+1})$．他

把霍普夫的定义做了稍微的修改,把 Σ 的 p 次同调群定义成因子群 $H_p(\Sigma) = C^p/R(K^{p+1})$.

拓扑概念赋予群结构很快导致了新研究方向.例如,与梅尔所定义相似,一个流形上定义的一组带有加法的微分形式可被看成一个阿贝尔群的复形.通过引进一个算子 d(外导数),把 k 维形式变为 $k+1$ 维形式,它与边缘算子有相同的特性,即是说 $d^2 = 0$.以同样的方式去定义闭链和边缘是可能的,于是便定义了**上同调群**,一般记为 $H^k(A)$.

在流形的同调群和上同调群的情形中,对空间指派群转换到指派空间之间的函数和它们相应的群.也就是,如果 $f: A \to B$ 是在两流形 A 和 B 之间的连续函数,此处把流形看作单纯复形,如果 $H_k(A)$,$H_k(B)$ 各自是 A 和 B 的 k 次同调群,那么就有一个定义好的群同态 $H_k(f): H_k(A) \to H_k(B)$.事实上,$H_k(f)$ 是在 k 的链 $H_k(f)(V_0V_1\cdots V_k) = f(V_0)f(V_1)\cdots f(V_k)$ 上定义的.能证明根据同调群理论这种对应是有意义的.也就是,闭链对应闭链,边缘对应边缘,因此就能定义商群之间一个合适的同态.对上同调群,情况类似.尽管群同态 $H^k(f)$ 是从 $H^k(B)$ 到 $H^k(A)$ 的对应关系,f 仍是一个可微函数 $f: A \to B$.

由于考虑定义在一个对象集合的成员上函数之间的关系以及一个不同集合的相关成员上定义的新函数,艾伦伯格(Samuel Eilenberg)(1913—)和麦克莱恩(Saunders Mac Lane)(1909—)得出了更加抽象的代数结构,即范畴.他们在 1945 年的一篇文章中,在某种意义上对克莱因的爱尔兰根纲领进行了推广.他们认识到,只要新的数学对象被定义,这些对象之间的映射的定义总是可以给出的.这样一个**范畴** $\mathscr{C}$就定义为一个二元集$\{A, \alpha\}$,其中 A 是抽象元素的集(例如群等),称为范畴的**对象**,抽象元素 α(例如同态)成为范畴的**映射**."[21] 这些映射满足一定的公理,包括适当的乘积映射的存在性.其中乘积映射满足结合律,并对每个 A 都有恒等映射.范畴的例子除开群和同态以外,还有拓扑空间和连续映射、集合和函数,以及向量空间和线性变换等.

按照他们的构想,艾伦伯格和麦克莱恩进一步引进了函子的概念,所谓的函子是范畴之间的函数.即如果 $\mathscr{C} = \{A, \alpha\}$ 和 $\mathscr{D} = \{B, \beta\}$ 是两个范畴.从 $\mathscr{C}$ 到 $\mathscr{D}$(协变的)**函子** T 是一对函数,两个都由相同字母 T 命名的,一个是目标函数;另一个是映射函数.目标函数把 $\mathscr{C}$中的每个 A 映成 $\mathscr{D}$ 中的目标 $T(A)$,而映射函数则把 $\mathscr{C}$中的每个映射 $\alpha: A \to A'$ 映成D 中的映射 $T(\alpha): T(A) \to T(A')$.这对函数必须把恒等映射映成恒等映射,而只要在 $\mathscr{C}$中存在 $\alpha\alpha'$,还要满足 $T(\alpha\alpha') = T(\alpha)T(\alpha')$.(对一个逆变函子,映射是倒置的,即 $T(\alpha): T(A') \to T(A)$ 和 $T(\alpha\alpha') = T(\alpha')\,T(\alpha)$.)例如,同调是从流形和连续变换组成的范畴到阿贝尔群和同态组成的范畴的一个协变函子,而对有限维向量空间 V 指派所有在 V 上的实值线性函数的向量空间 $T(V)$ 诱导了一个从向量空间和线性变换的组成范畴到本身的逆变函子.

令人惊奇的是,范畴和函子的研究曾被认为是为了抽象而抽象,但它们已被证明不仅仅在代数,而且在微分和代数几何近来的许多发展中都是非常重要的.其它在现代代数的似乎较抽象的发展中,包括近来的有限单群的完全分类,在其它领域也呈现出它的重要性.所以,虽然为抽象而抽象工作的数学家们常挨批评,但现在看来,就像已经过去的几个世纪一样,因为数学本身需要而发展起来的数学思想,常常被证明在解决现实世界问题时是非常关键的.

18.4　计算机及其应用

在 20 世纪晚期,当一个没有受过专门教育的人想到数学时,他首先想到的便是计算机.数学家们自己正逐渐地接受机器计算进入他们的学科.对许多数学家而言,笔和纸仍是最重要的工具.然

而自从20世纪50年代以来由于计算机的计算能力飞速发展,计算机已经进入了数学主流问题的研究中.现在,越来越多的数学家不仅使用计算机引出例子,而且还用来证明定理.令人感兴趣的是,许多多年沉寂的理论数学分支,由于在计算机科学领域的广泛应用已重新获得了人们的青睐.尽管本文没有篇幅对计算机的发展历史给出一个详细的描述,但我们将对其发展的重要事件和与它密切联系的几个数学分支作一简单的描述.

18.4.1　计算机的历史背景

当托勒密(Ptolemy)被迫使用大量的人力'计算器'去处理那些在他的《大成》中的许多表格时,机械化计算的梦想一定已经发生在古希腊时代了.在中世纪,一些伊斯兰科学家事实上的确使用了一定的工具去帮助他们进行计算,特别是那些与天文学有关的计算.到17世纪早期,在欧洲天文表格的计算都还是重要的,并且对数的发明部分也是为了帮助这一点.在较短的时间里,两位英国人德拉梅恩(Richard Delamain)和奥特雷德(William Oughtred)(1574—1660)以计算尺的形式独立地创造了对数表的物理形式,该计算尺即是一个按圆周移动的数字比例尺(后来改为按直线移动),数的乘法和除法,还有涉及到三角函数的计算,都能被容易地完成.

几乎同时,蒂宾根(Tübingen)大学的天文学和数学教授希克德(Wilhelm Schickard)(1592—1635)设计和制造了一台能自动完成加法和减法的机器,而且半自动地完成乘法和除法(如图18.6).希克德在1623和1624年给开普勒的信中描述了此机器,他还准备送给开普勒一台.但是这台为开普勒使用而制造的机器在它尚未完工时,于一场大火中被毁掉了.这台机器的残余副本和其设计者,在三十年的战争后根本就不存在了.因此希克德的发明装置对以后的工作没有任何的影响.大约20年以后,帕斯卡构造了一台处理加法和减法的机械装置(如图18.7),而1671年莱布尼茨构造了一台还能处理乘法和除法的机器.莱布尼茨相当确信,它的机器会有实际的伟大用处:



图 18.6　希克德的计算机.

> 我们可以说,它对所有那些从事计算的人来说都是梦寐以求的,大家都知道,如财务经理、房产管理人员、机修工、外科医生、地理学家、海军、天文学家及任何使用数学的行业

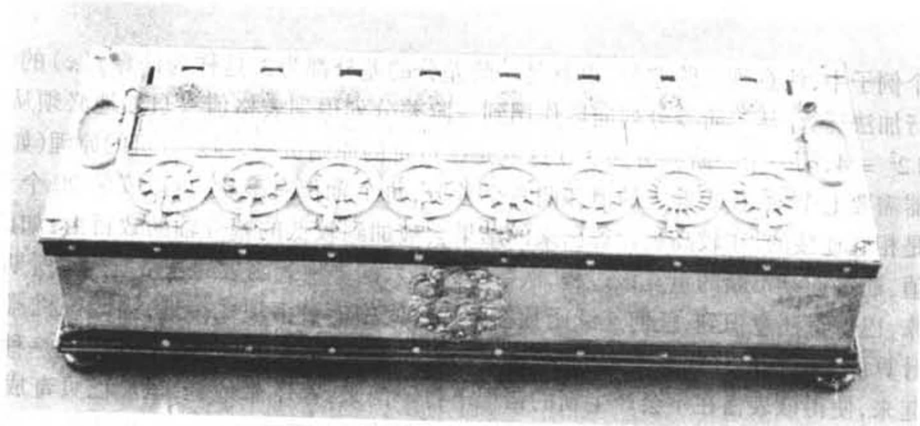


图 18.7　帕斯卡的机械计算装置的一个复制品(来源:Donges Neuhart 设计家公司).

中的工作者.但是就限制在科学的使用上而言,古老的几何学和天文学表格可能要被纠正,通过它的帮助,我们可以测量出所有的曲线和图形从而可以创造新的表格……更进一步……,它容易让任何人去创建一个自己的表格,以使他能够不费力气而且非常精确地进行他的调查研究……还有,使天文学家非常确信,继续锻炼计算上要求的耐心是不必要的……因为优秀的人士像奴隶一样花时间在计算上是不值得的,若机器被使用,其它任何人员都可以代作该计算.[23]

18.4.2　巴贝奇的差分机和分析机

很不幸,莱布尼茨的机器和其他人设计的改造模型在接下来的一个半世纪里在某种程度上都没有能以莱布尼茨预想的方式被实际利用.可能是因为手工运行机不能提供速度上的好处,数学工作者们仍继续使用手算.对于许多复杂的计算,特别地在对数和三角函数中,自然而然,表格被使用了,即使是表格最初还是通过手算的,仍频繁出错.直到工业革命在英国蓬勃发展和蒸汽机的发明,一个光辉人物,巴贝奇(Charles Babbage)大约在 1821 年设想使用这个新技术的思想去获得机器,它能提高速度和数学计算的正确率(图 18.8).



图 18.8　纪念巴贝奇和他的"计算机"的英国邮票.

巴贝奇认识到,n 次多项式函数值的计算可以通过使用 n 阶差分总是常数这样的事实而受到影响.举一个简单的例子,考虑下面关于函数 $f(x) = x^2$ 的表格:

x	$f(x)$	一阶差分	二阶差分
1	1		
2	4	3	
3	9	5	2
4	16	7	2
5	25	9	2
6	36	11	2

在这个例子中,注意到二阶差分,也就是一阶差分的差分都为 2.这样为计算 $f(x)$ 的值,惟一需要的是进行加法运算,从二阶差分列向后作用到一阶差分列得到表格值.(自然地必须从确定的值开始,例如 $2^2 = 4$,第一个一阶差分为 3.)这就是巴贝奇的原始机器,即差分机的原理(如图 18.9).制作该机器需要七个轴,代表着表的值和前六个差分,每个轴都有滚轮,可以安装 20 个十进制数,而这些轴是相互连接的.在较高级计算出来的结果会被加到较低的差分轴的数目上,如此等等,直到得到表值.通过连续不断的重复该过程,次数不超过六的多项式函数的表格就能按所期望的变量值计算出来.巴贝奇还意识到,任何连续函数在合适区间中能被多项式接近,因此该机器能被用来为科学家计算那个时代的任何有兴趣的函数的表格值.事实上,他的目的是将机器与一种打印盘的装置联系起来,使得该表格在不会产生新的错误下打印出来.很不幸地是,虽然巴贝奇成功地说服了英国政府在经济上支持他建造该差分机,但由于在制造精密部件方面的各种困难,由于最终政府对该计划失去了兴趣,还因为他本身又对用途更广的分析机产生了兴趣,完整的差分机模型始终未

能建造出来.

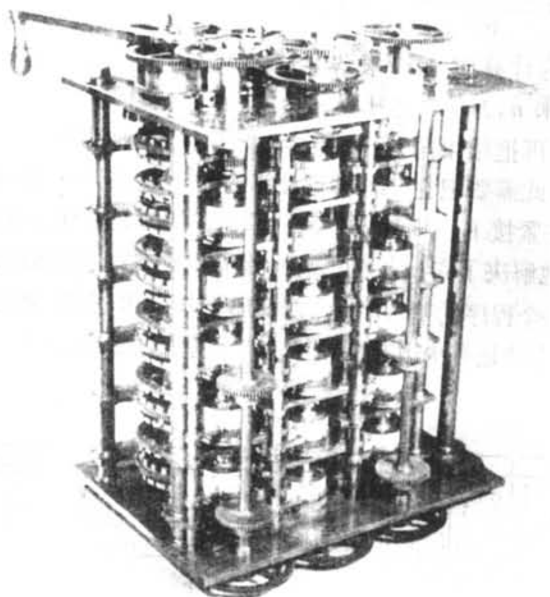


图 18.9 巴贝奇的差分机的现代模型(来源:Donges Neuhart 设计家公司).

巴贝奇在1833年开始了他的新项目,到1838年,已经详细拟定出了基本设计方案.他的新机器含有许多今天计算机的特征.在轴轮上构造了许多齿轮作为硬件.该机器包含两个基本部分:**储存部分**和**加工部分**.储存部分是一个保持数值变量的地方;直到变量被加工,并保存其结果.而加工部分是一个进行大量运算的地方.为控制运算,巴贝奇采用了雅克(Joseph Jacquard)(1752—1834)(图18.10)的思想.雅克通过一种描述织布机理想模式的穿孔卡的引进为法国的纺织工业实现了自动化.巴贝奇就这样设计出了自己的穿孔卡系统,该系统能容纳数字值和对该机器的指令.虽然巴贝奇从未把他的分析机完整描述成文,而且实际上从未得到经济上的援助,但它的确为后人留下了300大张的工程图纸,每一张图纸大约2×3英尺,以及关于他的思想的成千上万页的细节描述.现代学者通过研究这些论文总结出,那个时代的技术可能足以建造一台这样的机器,但部分因为英国政府没有足够的兴趣去为如此一项大的项目施以财政援助,此机仍是一个理论上的构造.



图 18.10 法国邮票上的雅克.

1840年,巴贝奇为聚集在都灵的一群意大利科学家做了一系列有关分析机的实用性的研讨会,他们中的一位总结了这次研讨会,并写成了论文发表.这篇17页的论文在1843年被翻译成英语,且由洛夫莱斯伯爵夫人,艾达·拜伦·金(Ada Byron King)(1815—1852)用额外的40页注释做了补充.在她的注释中,洛夫莱斯不仅在关于发动机的细节功用上在文章的许多部分做了展开,而且对怎样处理特殊问题做了明确的描述,这样,在出版物中她第一个描述了今天被我们称作的计算机程序.她以计算一个伯努利大数的程序为例.她首先从描述伯努利数为下列展开式的系数 B_i 开始

$$\frac{x}{e^x-1}=1-\frac{x}{2}+B_2\frac{x^2}{2!}+B_4\frac{x^4}{4!}+B_6\frac{x^6}{6!}+\cdots\quad(\text{见第 14 章习题 3}).$$

通过使用 e^x 的幂级数展开的代数方法,洛夫莱斯以下面的形式重新写了这个方程式

$$0 = -\frac{1}{2}\frac{2n-1}{2n+1} + B_2\left(\frac{2n}{2!}\right) + B_4\left(\frac{2n(2n-1)(2n-2)}{4!}\right)$$
$$+ B_6\left(\frac{2n(2n-1)\cdots(2n-4)}{6!}\right) + \cdots + B_{2n},$$

而 B_i 都可以递归地计算出来. 这样为计算 B_{2n}, 需要 3 个数字 1,2,n 以及已被计算出来的数值 $B_i(i < 2n)$. 然后需要指令卡用 2 去乘 n, 并从结果中减 1, 和在结果上加 1, 二者所得的结果再相除, 再与 $-1/2$ 相乘, 然后将 $2n$ 除 2, 再把结果与 B_2 相乘, 如此下去. 这种计算的某些结果, 比如 $2n-1$, 在计算中被使用了好几次, 因此需要把它调到进行计算的数字记录器中. 在计算的过程中, 机器被指示从 $2n$ 中减去一个整数, 在紧接下一步就依靠其结果是正的或是 0 做出决定. 若是 0, 对 B_{2n} 的方程式就完成了, 机器很容易地解决了它; 若是正数, 机器还要重复许多以前的步骤. 不难看出在洛夫莱斯的描述中已有了一些现今程序的基本概念, 如循环与判定步骤等. 而且, 她已画好了附以注释的以上程序的详细图表, 也许这是有史以来的第一个"流程图"(图 18.11).

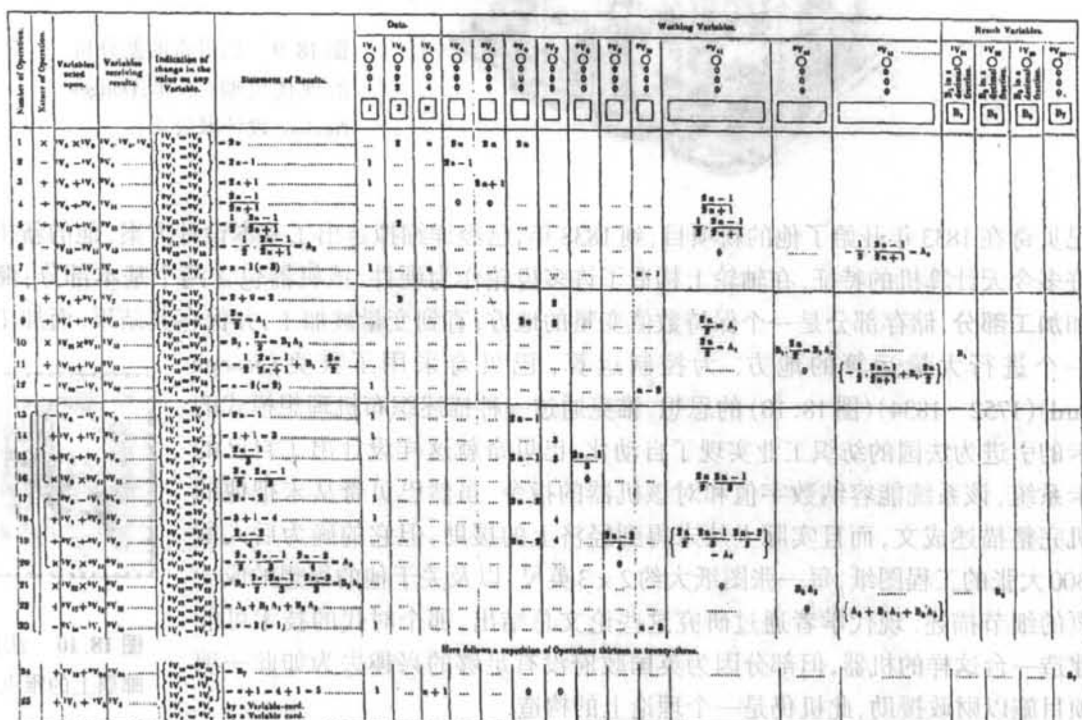


图 18.11　艾达 · 金的计算伯努利数的流程图表.

除了讨论分析机的基本功能外, 洛夫莱斯还描述了它能做的所有工作, 并详细指出它能完成符号代数运算, 还有算术运算. 但是, 她指出:

> 解析机丝毫没有自称可以**创造**出什么东西. 它能做一切**我们知道如何命令它做**的任何事, 它**遵循**分析学. 但它没有**参与任何**分析关系或事实的能力. 它的作用是帮助我们**利用**所熟悉的东西, 但它可能以另外的方式对科学施以**间接**式的相互的影响. 因为, 如此分配与组合事实和分析学的公式, 使得它们能够变得最容易和最快速地顺从于分析机的机械组合, 从而科学中的许多学科的关系和性质必定得到了更清楚的了解, 并得到更深刻的

研究.十分显然,根据一般原则,在对数学事实设计一个能够记录并能付诸实用的新形式时,所能得到的评论大概是说,还应该在学科的更为理论的方面有所作为.

今天关于计算机对数学发展的局限性和含意也不可能写出比这更好的描述来了.

人物小传

艾达·拜伦·金(洛夫莱斯伯爵夫人(1815—1852)(Ada Byron King, Countess of Lovelace))

洛夫莱斯伯爵夫人是拜伦六世勋爵乔治·戈登的孩子.在她出生后五个星期,他离开了英国,以后,他再也没见到她.她是由母亲养大的,她的母亲叫米尔班克,是一个学数学的学生.因此,洛夫莱斯比那个时代的一般女孩获得了更多的数学教育.虽然她从未上过大学,但她被私人教导,且能向著名数学家讨教,其中包括弗雷德和德·摩根.1833年,她遇到了巴贝奇,并很快对他的差分机发生了兴趣.她的丈夫,洛夫莱斯,1840年成为了皇家学会的会员.通过这个关系,洛夫莱斯获得了她所要的书籍和论文的途径,以继续她的数学研究.她的主要数学工作,即本文讨论的,是意大利数学家梅纳布里(L.F.Menabrea)一篇论文的翻译和详细注释.梅纳布里研究了巴贝奇的分析机.使人感兴趣的是,该论文仅以她姓名的起首字母A.A.L出版.明显地,在19世纪中期的英国对妇女去出版数学著作被认为是不适当的.

18.4.3 图灵和可计算性

在19世纪中期的英国,巴贝奇的有关实际的分析机的思想没有被变成现实的原因之一,是因为没有任何可见的社会需求,且不得不去说明该构造花费巨大资源的正确性.在该世纪巴贝奇的设计方案后,虽然许多计算装置和类似计算机的装置被设计出来,但一般都是和解决具体的需要大量手工计算的数学问题有关.由于两次世界大战,特别是第二次世界大战中军事的需要,导致了第一台电子计算机的实际建造.另外,其它的理论思想在二次世界大战前几年就已产生了,其中之一是图灵(1912—1954)工作中的可计算性思想.这些都是计算机发展的基础.

人物小传

阿兰·图灵(1912—1954)(Alan Turing)

图灵的父亲是英国驻印度办事处的一位官员,他认为他的儿子应该在英格兰长大成人.所以图灵的父母在他的成长岁月里很少与他相聚.图灵于1931年进入剑桥大学的国王学院学习数学,并在4年之后取得硕士学位,其毕业论文是关于高斯误差函数的.然而,在这之后不久,他又开始热心致力于一个新问题——希尔伯特的判定问题,并且发表了一篇论文.在这篇论文中他创立了图灵机器的概念.几乎与此同时,普林斯顿大学的丘尔奇(Alonzo Church)发表了对同一问题的另外一种解法,于是图灵决定去普林斯顿与丘尔奇合作.1938年他返回英格兰国王学院.这时爆发了二战,他应召到白金汉郡的布勒奇利公园的政府密码部门服役.在接下来的几年里,正是在那里图灵成功地破解了德国的"谜"密码,这最终成为挫败纳粹德国的重要一环.

战后,图灵继续致力于研制自动计算机,并加入了国家物理实验室参加研制计算机.1948年后他在曼彻斯特大学继续这项工作.但是,图灵的颇有前途的事业令人痛心地终止了.1953年他因"公开行为粗鄙"而被捕.事实上,他是一名同性恋者,在那个时期的英格兰,公开的同性恋行为是违法的.对他的惩罚是接受心理分析和一种设计来"治愈"这种疾病的荷尔蒙治疗.不幸的是,治疗比他的病更糟糕.并且,在一段时期的消沉中,他吃了一只含氰化物苹果于1954年6月自杀死去.

图灵对可计算性问题感兴趣.所谓的可计算性问题,就是对计算是什么和一个所给定的算法是否能被实际执行等问题给出一个合理且精确的问答.为回答这些问题,图灵从一般的计算过程中抽取了必要的部分并且根据理论机器的方式把它系统阐述出来,这就是现今闻名遐迩的图灵机.而且他还证明了存在"通用"的图灵机,只要给出合适的指令,它就能对任何特殊机器可以进行计算的数和函数进行计算.

图灵机的思想呈现在 1936 年的一篇文章中,主要由三个基本的概念构成:一个状态的有限集,或称组态$\{q_1, q_2, \cdots, q_k\}$;一个符号有限集$\{a_0, a_1, a_2, \cdots, a_n\}$,这些符号能够被机器读和写(其中$a_0$是一个空白符号);状态改变和符号被读的过程.机器为了完成工作,必须通过运行在它上面的(无穷)纸带使机器获得一些指令.而这个纸带被分作方格,其中有限个方格存有非空符号.在任何给定的时间,机器上都有惟一的一个方格,设是第 r 个,存有一个符号 S_r.为把问题进一步简化,符号给出的可能指令局限于让一个新的符号代替方格上的符号,即把纸带向左或向右移动一格来改变机器的状态.这样在任何给定的时刻,通过特别的函数关系的定义,偶对(q_r, S_r)就将决定机器的行为.如果函数没有定义在特别的(q_r, S_r)上,机器就会停止运行.而正是机器打印出来的符号,或至少其中一部分代表了计算出来的数.他的论文中许多支持图灵的论点的论证在我们刚才所描述的运算中对计算一个数来讲是必须的.

例如,图灵构造了一台机器来计算序列 010101….这台机器有 4 个状态 q_1, q_2, q_3, q_4,并能够打印两个符号,0 和 1.最初机器的纸带是完全空白的,并且机器从状态 q_1 开始.机器所用的指令如下:

1. 如果它处在状态 1 并且读到一空白的方块,则打印 0,向右移一格,转换到状态 q_2;
2. 如果它处于状态 q_2 且读到一空白的方块,则向右移一格,转换到状态 q_3;
3. 如果它处于状态 q_3 且读到一空白的方块,则打印 1,向右移一格,转换到状态 q_4;
4. 如果它处于状态 q_4 且读到一空白的方块,则向右移一格,转换到状态 q_1.

尽管因为技术上的原因,图灵将数字安排成间隔排列的样子,但还是很容易看出,此机器确实能够实现设计的功能(如果数字是连续打印的,则此机器还可以简化).这个例子还没有表现出来的一点是:让纸带向两个方向都能移动的原因是赋予机器记忆.这样,机器可以重复读一特定的方块,并根据当时的状态来做出不同的反应.机器凭着这种方式可以"记住"以前写下的数字并在以后的计算中使用它们.

正是记忆的可能性产生了图灵的论文中最惊人的部分 —— 证明一台可计算任何可以计算的数字的机器的存在性.对此,图灵的想法是先得到对任意给定机器的指令集,就像上述的指令一样,然后把这个指令集系统转换成一系列符号 —— 机器的标准描述.然后,提供这台全能的机器一张含有标准描述的纸带,紧接着标准描述的是最初输入给机器的符号.事实上,图灵能够以前面提到的函数关系的形式对这台机器的所作所为给出一个更清晰的描述,其主要思想是机器循环工作.每个循环表示:首先,浏览一特定机器的标准描述;其次,读该机器输入的一个方块;最后做出相应的反应.尽管,图灵没有设想出一台像他所证明存在的机器的物理构造,但他的想法直接导致了可以设计来做任何计算的"全能"计算机的产生.当然,对于机器的大小,和程序的长短还有一些物理上的限制.而这些限制在图灵的理论模型中并不存在.然而,现代技术如此频繁地减弱这些限制,以致于计算机正逐步向图灵的"全能"计算机转变.

18.4.4 香农和开关代数

数学思想在计算机的构造中另一更为直接的应用是 1938 年香农(Claude Shannon)(1916—)提

出的，这是他在麻省理工学院硕士论文中的一部分.在他的论文中，香农将1世纪前布尔发展起来的逻辑代数应用于构造具有某些性质的开关线路中.事实上，这些电路正是计算机内部结构的基础.香农认识到了每个电路都可以用一个方程组来表示，并且对方程组求解所需的工具恰好就是布尔逻辑代数.这样，给定一个要构造电路的特征，用布尔逻辑代数就能把方程组转换成最简单的形式.从这些形式出发立刻就可得到电路的结构，也可以用这种方法来分析电路，只需将布尔逻辑代数应用于复杂电路的方程组，然后将它化为更简单的形式.

香农从简单的只有开或关两种状态的开关入手，1代表"开"，0代表"关"，把两个开关串联用布尔运算"+"表示，并联用"·"表示(如图18.12).香农用相应的电路解释对这两种运算给出了以下的假设，而这些假设的正确性使得用布尔代数分析开关电路成为可能.

1. $0\cdot 0=0$；$1+1=1$，两个关电路的并联是关，而两个开电路的串联是开.
2. $1+0=0+1=1$；$0\cdot 1=1\cdot 0=0$，一开电路与关电路的串联是开，而一关电路与一开电路的并联是关.
3. $0+0=0$；$1\cdot 1=1$，两关电路的串联是关，而一开电路与一开电路的并联是开.

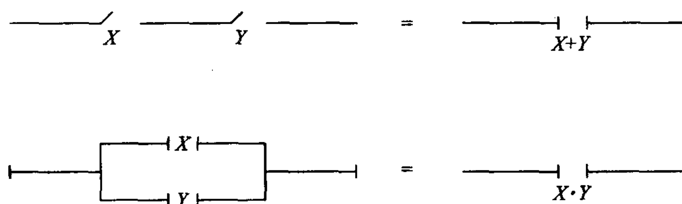


图 18.12　串并联开关.

电路中 X 表示开关，香农说明 X 只能取0和1两个值.这样布尔代数中的法则，包括两个交换律、两个结合律、乘法对加法的分配律、加法对乘法的分配律都可以用检验每种可能的情形来验证.他还引入了 X 的负值，记为 X'.当 $X=0$ 时，$X'=1$；当 $X=1$ 时，$X'=0$.他还验证了一些加法法则，例如，$X+X'=1$，$X\cdot X'=0$，$(X+Y)'=X'\cdot Y'$，$(X\cdot Y)'=X'+Y'$.香农利用布尔对函数的展开式和交换乘法和加法的方法将其对偶.例如：

$$f(X,Y,Z,\cdots)=[f(0,Y,Z,\cdots)+X][f(1,Y,Z,\cdots)+X'].$$

将两边都加 X，就可以建立一个非常有用的法则(注意分配律)：

$$X+f(X,Y,Z,\cdots)=X+f(0,Y,Z,\cdots).$$

有了这些为开关电路建立起来的布尔法则，香农就可以分析并合成电路.例如，香农对电路(图18.13)给出了它的代数表示：

$$W+W'(X+Y)+(X+Z)(S+W'+Z)(Z'+Y+S'V).$$

利用布尔代数的各种法则，包括前面段落里使用过3次的特殊法则，他将上式首先简化为

$$W+X+Y+(X+Z)(S+1+Z)(Z'+Y+S'V),$$

然后化成

$$W+X+Y+Z(Z'+S'V),$$

最后化成

$$W+X+Y+ZS'V.$$

后面的式子具有比最初的式子更简单的电路表示.

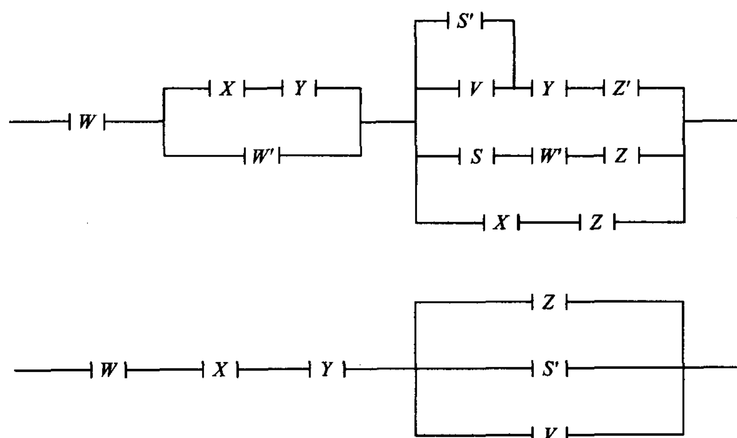


图 18.13　利用布尔代数简化电路.

作为一个合成具有给定特征的电路的例子，香农说明了如何去构造一个数，而这个数是两个给定二进制表示数的和.设有两个数用 $a_na_{n-1}\cdots a_0$ 和 $b_nb_{n-1}\cdots b_0$ 表示.并且它们的和用 $s_{n+1}s_n\cdots s_1s_0$ 表示,如果 $a_0=1$ 且 $b_0=0$ 或者 $a_0=1$ 且 $b_0=1$,则 $s_0=1$,否则 $s_0=0$.除此之外还有一个进位项 c_1,如果 a_0 和 b_0 都是 1,则 $c_1=1$;否则 $c_1=0$.这样,s_0 用方程 $s_0=a_0b_0'+a_0'b_0$ 来表示,而 c_1 用 $c_1=a_0b_0$ 表示.每一个 $s_j,j\geqslant 1$,不仅需要 a_j 和 b_j 的加法,而且也需要进位 c_j.所以 s_j 的表达式是 $s_j=(a_jb_j'+a_j'b_j)c_j'+(a_jb_j'+a_j'b_j)'c_j$,而且下一个进位项 c_{j+1} 为 $c_{j+1}=a_jb_j+c_j(a_jb_j'+a_j'b_j)$.由这些方程决定的加法器的电路构造是现代计算器和计算机中加法设计的基础.

18.4.5　冯·诺伊曼的计算机

图灵和香农的工作只是实现现代计算机所必须要解决的理论和应用问题的两个方面.有很多人都致力于这些问题,特别是在 20 世纪 40 年代.但是对最终结果作出重大贡献的人是冯·诺伊曼(Von Neumann)(1903—1957).他在二战后迅速在普林斯顿高等研究院召集一群杰出的科学家和工程师组成了一个小组,他们的任务是掌握两种早期计算机——ENIAC 和 EDVAC 在战争期间的发展情况,并且将其与现有的理论知识相结合.他们的目的是开发一种机器,这种机器被小组中一位科学家称为“目前已有的最复杂的研究工具……科学家们对于研制出这种工具的可能性表现出极大的兴趣.并且它的构造可解决目前人们只能梦想的东西.”[26]

冯·诺伊曼领导下的小组决定将计算机分成四个部分——计算部分、贮存部分、控制部分和输入输出部分.前两部分分别相当于巴贝奇的加工部分和储存部分.计算器——现在普遍称为中央处理器——是执行基本运算的地方,这些运算不能被简化.这些基本运算都被存入机器,例如前面提到的加法,而其它的运算都可以通过指令由这些基本运算构造出来.我们知道巴贝奇的分析机器的数字系统是十进制的.但随着电子——而不是机械的——表示数字的仪器的发展,用二进制表示数字更为简单.所以任何一种存贮数字的仪器都有两种状态,开或关来代表 1 和 0 两种可能性.事实上,冯·诺伊曼设计了一系列能够有效的将十进制转化成二进制和将二进制转化成十进制的仪

器,操作者可以输入普通的十进制数字并得到二进制的结果,而不考虑速度和构造机器的复杂性.

机器的存贮器要承担两种不同的任务,存贮在计算机中要用到的数字和存贮如何进行计算的指令.但是,由于指令本身也可以以合适的数字形的码存贮起来,所以机器只需分清什么是真正的数字,什么是转换成码的指令.进一步的,为了要在使用者需要的"无限"存贮和工程师只能构造的有限存贮之间达到平衡,研究小组决定将存贮器设计成阶梯式的.这样,少量的存贮单元可以马上使用,而更多的存贮单元则以较慢的速度加以利用.他们还决定,为了要在一合理的物理空间上获得一足够的的存贮空间,存贮一位数字的单元必须是某一大块存贮单元中非常小的一部分.

控制器是放置指令的部分,也就是说它是存放着实际要遵守的命令的地方.再者,设计人员必须作出平衡,一方面是设备要简单;另一方面是要想使机器有效工作就得有大量的不同指令.无论如何,控制程序中有许多更为重要的方面,其中之一是机器要有重复使用一系列指令的能力.这一点洛夫莱斯女士也意识到了,但是因为机器还不知道什么时候结束重复使用某些指令,这就需要设计出一种指令来让它知道.更进一步的,控制单元需有一指令将输入输出和机器结合成一个整体.事实上,冯·诺伊曼对输入输出兼备打印和图表输出有特殊的兴趣,因为他认识到,一些更为重要的计算结果最好以图表的形式来研究.

基于冯·诺伊曼的设计思想的计算机最终在高等研究院设计出来,并在1951年完成,成为以后设计出来的更先进的计算机的模型.20世纪后半叶,计算机方面的技术成就不仅增强了计算机的功能,而且减小了计算机的体积,这些恐怕是20世纪40年代的小组成员所不能想象的.现在,计算机已成为日常生活中重要的一部分,没有它我们就连最平常的工作也做不了.尽管我们不能涵盖计算机应用中的每一个方面,但我们在本章最后将简单介绍几个数学分支,这些分支对于现在计算机的普及发挥过作用,而且也曾受过计算机的深刻影响.

人物小传

冯·诺伊曼(1903—1957)(John Von Neumann)

冯·诺伊曼(图18.14)生于布达佩斯一个富有的犹太家庭,在布达佩斯大学取得博士学位.1930年受邀进入普林斯顿,在此之前他在柏林和汉堡教书.三年后,他被选为高等研究院的特许成员,他此后一直担任这个职位.冯·诺伊曼对应用和理论研究都很擅长,这是最后的这类数学家中的一个.每年他在这两个领域都有数量稳定的论文发表.在理论数学方面,他特别擅长分析和组合.他有超凡的能力能透过问题复杂的表面提取其精髓,使之可用数学方法分析.他在应用数学方面的才华使他在二战期间倍受重视,在发明现代计算机后很快就获了奖.1954年,他加入了原子能委员会,直至1957年不幸死于癌症.



图18.14 匈牙利邮票上的冯·诺伊曼.

18.4.6 查错与纠错

当将大量信息从一个地方以电子方式传送到另一地方时,我们通常希望在接收的时候出的错越少越好.正如第一篇有关查错与纠错的论文作者海明(Richard Hamming)(1915—)所指出的那样,这种能力非常重要,当某个操作需时很长,而且是在一个监控设备很少的系统中进行的,在这种系统中一个可能的错误就可能毁掉整个操作,并且传输过程中可能会有噪声,例如,发生"堵塞".

1947年海明对编码产生兴趣是因为他使用早期的贝尔系统,这种系统在周末是依赖于计算机的.周末的时候,机器无人操作,在此情况下机器会跳过某些它认为有错误的问题而直接进入下一

个问题.海明意识到,要是让计算机不仅可以查错而且还可以纠错的话倒是一件有价值的事,那样他就可以完成工作了.1948 年他继续研究他的纠错码,但是由于专利审查,直到 1950 年他才能把它们公布于众.

在他的论文中,海明假设每一个码字均由 n 位二进制数组成,其中 m 位携带信息而 $k = n - m$ 位数字用来查错和纠错.他的目的是以尽可能少的检查数来确定一个单个错误的纠错码.每一个检查数都是对码字的"对等" 检查,也就是说,它等于 0 或 1,根据要查的码中 1 的总数是偶还是奇.想法就是,在接受一个给定的码字时,其中由 1 和 0 组成的检查码都会被读作一个二进制数,并会被要求给出任一单个错误的位置.因此,k 位检查数必须能够描述 $m + k + 1 = n + 1$ 种可能性,因为错误可能出现在 n 位中的任何一位,也有可能根本就没有错误.这决定了 k 位的字符串可能的数目,即 2 的 k 次幂至少要与 $n + 1$ 相等.换一种说法,m 与 n 的关系由下面的不等式给出:

$$2^m \leqslant \frac{2^n}{n+1}.$$

例如,一个含有 7 个位置的码字要有 3 位作为检查码而 4 位用来携带真正的信息.即一个 7 位长的字有 $2^7 = 128$ 种可能,其中只有 $2^4 = 16$ 个可能用于真正的码.

海明为他的方法发明了一种几何模型,将 n 位的码字看成为两个元素的域上的 n 维空间中的单位立方体的一个顶点.他进一步引进了这种空间中的度量,使得这样的两个顶点之间的距离 $D(x,y)$ 就是它们之间不同坐标的数目.例如,如果 $n = 3, x = 001, y = 111$,它们之间的距离 $D(x,y) = 2$.海明指出,如果任意码点之间的距离都不小于 2,那么任何一个单个错误都可以被查到,因为这样的错误会将一个码点变为另一个码点,而这个码点与立方体中某一个码点之间的距离是 1.如果任意码点之间的距离是 3,那么任何一个单个错误都会使新的码点与某一真正的码点之间的距离变短,这样错误也是可查的.结果是确定纠错码的问题与在一个空间中找到一个点的子集的问题相同,使这个子集的点在空间中的距离至少是一个最小值.特别地,如果这个最小值是 3,则每个点都可被一直径为 1 的球围起来,使得任何两个球都没有公共点.(以 x 为中心,半径为 r 的球是与 x 距离小于或等于 r 的点的集合.)现在,海明可以利用他的几何模型来回答有多少真正的码的问题了.由于每个半径为 1 的球的表面上都有 n 个点,算上中心点共有 $n + 1$ 个点.又因为整个空间共有 2^n 个点,所以最多有 $\frac{2^n}{n+1}$ 个球,所以码点数 2^m 不会大过这个数,这正如前面提到的那样.

在他的论文中,海明特别给出了一个 7 位长的码字,它满足检查字的条件,这个码就是现在众所周知的海明码.当时海明隐约感觉到,而事实也证实了:这个码中 4 位的实际码组成的集合是一个群,更确切的说是一个 7 维空间的 4 维向量子空间,这个 7 维空间是由所有 7 位长的二进制字符串组成.最终,海明码可以被描述成适当向量空间上某个线性变换的核.由海明开创的码的研究引出了许多其它的数学分支,有关于这方面的内容读者可在近期出版的书中读到.[27]

18.4.7　线性规划

线性规划处理的是线性函数 $a_1x_1 + a_2x_2 + \cdots + a_nx_n$ 在一定的约束条件下达到最大或最小的问题,其约束条件是变量 x_i 的线性不等式.奇怪的是,尽管对线性方程组的解的研究已有了 2000 多年,但是在二战之前关于线性不等式的研究却很少,而关于线性函数最大值的研究就更少了.只有傅里叶(Fourier) 在 1826 年考虑过这样的问题,但是也没有进行更深入的研究.

近代对线性规划的研究有两个起源 —— 军事和经济.俄罗斯的康托洛维奇(Leonid V. Kantorovich)(1912—1986) 是研究经济问题的数学家之一.他在 1939 年写了一本书,名叫《生产组织

和计划中的数学方法》.康托洛维奇相信工厂或者在整个工业组织中提高生产力的一个方法是对单个机器的工作的分配,不同供应商的定货,原材料的种类,燃料的种类,等等进行改进.康托洛维奇是第一个意识到这类问题都可以用相同的数学语言来表达并且可用数值方法来求解的人.然而,由于种种原因,康托洛维奇的研究工作并没有被苏联的经济学家或数学家继承下来,所以对他研究过的这类问题的一般数学解却是首先在美国发现.

正是在二战期间应美国空军的要求,线性规划问题才在美国得以研究.他们提出的是这类问题:安排特定的部门为战争服务,负责技术人员的培训,设备的供应和维修等.很显然对这类问题的各个方面的有效调剂要求新的数学方法,而这些方法在 1947 年左右才发展起来.那时空军建立了一个名叫最优规划的科学计算工程(SCOOP) 的工作小组,丹齐克(George Dantzig)(1914—) 是这个小组的主要成员.正是他研究出了解决线性规划问题的单纯形法的基本思想.

首先,确定解的适定集合,即在一适当维数空间中确定一个凸多边形,它包含了这组线性不等式的所有解.接下来,沿着凸多边形的边从一个顶点移向另一个顶点使线性函数的值达到最大,这种方法开始被丹齐克认为效率低而摒弃不用,但最终发现它还是确定目标解最有效的方法 —— 总能在一个顶点上得到一个解.掌握了丹齐克的方法后,冯·诺伊曼很快意识到线性规划问题与博奕论之间的关系,博奕论的基本理论是他在 1944 年研究出来的.冯·诺伊曼对线性规划问题和博奕论的数值解问题提出了多个建议.

尽管简单的线性规划问题可以用手算,但是令人们感兴趣的应用问题大多含有大量的变量和方程组,这样就需要某种机器来计算.由此,对一重要问题的单纯形法的使用于 1947 年秋季在早期的电脑上首次实验成功.接下来的几年里,几种计算方法被研究出来,使得新发展的计算机可以用来解决带有大量变量和方程组的线性规划问题.事实上,几十年来,线性规划的应用随着现代计算机的计算速度和计算能力的提高而迅速发展.[28]

18.4.8 图 论

现代术语中的图由一非空集合 V 和集合 E 组成,其中 V 的元素称为顶点,E 的元素称为边,每一条边连结两个顶点.在几何术语中,图的边就是连接两个顶点的弧.在西方,图论起源于欧拉 1736 年对哥尼斯堡七桥问题的解决.该问题在 14.3.7 节中曾讨论过.尽管欧拉自己用代数方法解决了此问题,但用点代表城市,边代表桥很容易就可以画出一个图来.另一个很有趣的后来成为图论的一部分的问题是由哈密顿于 1856 年建立起来的.事实上,哈密顿将问题转化成了一种游戏,并于 1859 年市场化.这二十点(Icosian)游戏由一个具有 20 个顶点的图组成,根据不同条件放置棋子.主要规则是一个子总是放到一条边的第二个顶点上,而这条边已被放置上一个棋子(图 18.15).哈密顿给出的第一组附加条件是给定放置在五个初始点的棋子,其余的子连续放置使得最后一个与最开始的一个相连.用更现代一点的说法来讲,哈密顿的问题就是要找到一个经过每个点一次且仅有一次的环路,与欧拉不

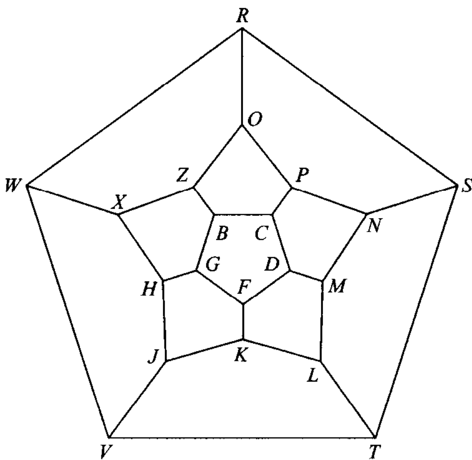


图 18.15 哈密顿的二十点游戏.

相同的是，欧拉的问题是要找到经每条边一次且仅一次的路径，哈密顿给出了几种解决他提出的问题的方法，但是他没有就更一般的情形确定这样的路径是否存在. 几乎与此同时，英国的一位牧师，寇克曼(Thomas P. Kirkman)(?—1892)，写了一篇简短的文章就更一般的情形讨论了相同类型的问题. 事实上寇克曼描述了一类不存在这种环路的图.

最早从纯数学的角度研究一类特殊的图开始于 1857 年凯莱的一篇文章. 受到对微分算子的可能的组合的研究的启发，凯莱定义并分析**树**的概念，树是一种不含圈的连通图，因而边的数目比点的数目少一. 特别地，凯莱研究了一种有一个特定的顶点作为根的树，即**根树**. 他说明了具有两个、三个和四个顶点(凯莱称他们为**结点**)的根树的可能的情形，或者等价地，具有一条、两条或三条边(以植物学的类比称为**树枝**)的根树. 依据一个巧妙的组合论证，凯莱为确定具有 r 个树枝的不同树的数目 A_r 建立了一递推公式(图 18.16). 例如 $A_1 = 1, A_2 = 2, A_3 = 4, A_4 = 9, A_5 = 20$. 1874 年凯莱将他的结果应用于化学的同分异构体的研究，几年之后他成功地建立起给定顶点数的无根树的计数公式.

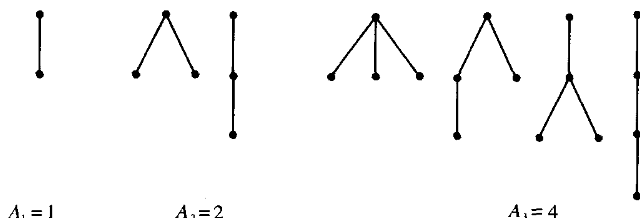


图 18.16　$r = 1,2,3$ 时 r 条树枝不同的树.

在近代图论发展中起着重要作用的问题是四色问题，这个问题从 1852 年提出起就一直被许多数学家寻求解决. 当年 10 月 23 日德·摩根在给哈密顿的信中描述了这个问题："今天我的一个学生古塞里(Frederick Guthrie)要求我就一个事实给予他回答，但这个问题我以前不知道，现在也不知道 —— 他说把一个图形分割成几块并且染上颜色，使得任何有公共边界的两块颜色都不同，只要用四种颜色，不会用更多的颜色. 我的学生说他猜测用这种方法就可以为英格兰地图着色. 我越想这个事实就越明显."[29] 德·摩根想不出要用五种颜色给地图着色的情形，尽管他认为四种颜色就够了是显然的，但是他却不能给出证明. 哈密顿对这个问题不感兴趣，但是在接下来的 20 年里，凯莱和皮尔斯(Charles S. Peirce)(1839—1914)花了很多时间寻找证明却没有成功. 1879 年，肯普(Alfred Kempe)(1849—1922)发表了一个被认为是正确的证明，第二年泰特发表了另外一种证明. 然而到 1890 年发现这两个证明都是错的. 希伍德(Percy Heawood)(1861—1955)指出了肯普的错误，但是作为他的结果的一部分，他证明了总是能够用五种颜色把地图着色.

20 世纪早期有很多人致力于这个问题的研究，但是建立起图论与地图着色问题之间的联系是惠特尼(Hassler Whitney)，这是最终证明四色问题的关键，并引起了后续的一系列研究. 1931 年，惠特尼定义了地图的对偶图，令图的顶点与地图的区域之间存在一一对应关系，如果两点之间有边连接，当且仅当对应区域之间有公共边界. 如果我们把图的着色定义为颜色对顶点的分配，使得同一条边连接的两个顶点颜色不同，那么地图的四色定理就等价于图的四色定理. 这样，就有可能将新发展起来的图论的一些结果应用在这个老的问题上.

正如本章开头提到过的，1976 年阿佩尔(1932—)和哈肯(1928—)对图论的四色定理给出了证明. 他们的证明方法是将定理演化成在一个图中可能出现的大量子图，然后研究这些子图着色的可

能性.为了要在可行的时间内完成这项工作,阿佩尔和哈肯不得不大量地使用计算机.也就是说执行单项检查的次数是如此之多,以至于没有任何人,或是任何一群人有可能用手计算出来.事实上,阿佩尔和哈肯在计算机上花了六个月左右的时间才完成这项工作.尽管这项定理的证明后来被简化了一些,但是不使用计算机的证明看起来似乎是不可能的.

大量依赖计算机的证明在数学界是全新的现象.计算机,自从它们问世以来,就用来帮助数学家们提出猜想,但是,四色定理的证明是计算机首次真正用来构造形式的证明.正如所预料的那样,这个证明自从它出现就引起了很多争论.很多数学家仍然不能接受这个证明的有效性,因为对于接受一个证明的一般标准是它已被数学界的许多数学家所检验.尽管计算机程序本身可以被检验,但是数学家们却不可能去检查计算机运行中的每个细节.这场对使用计算机的争论将如何结束还不得而知.其他重要的问题在计算机的帮助下得到解决也是有可能的,或许有一天四色定理也能用传统的方式证明出来,但是无论如何,一场新的关于是什么构成了数学证明的争论已经开始了.

最近对有限单群的分类又为这场争论添油加醋.因为这项分类有几百个数学家写了上千页纸,看起来没有任何人能够读懂并证实整个内容.然而,在这个特殊的事件中大多数数学家接受了证明是正确的,因为它的每一段都经得起传统证明方法的检验.但是,对于一个人怎样才能说服其他的人一项数学结果是正确的,问题还是存在.这个问题,自从人类有历史以来就有人在问,看样子将来还会这样问下去.

人物小传

哈斯勒尔·惠特尼(1907—1989)(Hassler Whitney)

惠特尼,生于纽约,在耶鲁大学获得物理学和音乐学士学位(1928,1929),并于1932年在哈佛大学取得数学博士学位.在普林斯顿大学做短暂的停留后他返回哈佛,在那里他执教直至1952年转到高等研究院.惠特尼主要的数学研究是微分拓扑,在这个领域中他做出了几个重要的定理,并开创了许多方法,所有这些都为这个学科奠定了基础.在他的非数学的爱好中有音乐——他是普林斯顿社区乐团的团长——和爬山——他在新罕布什尔的埃依山脉开辟了一条新路线.他发现了美国数学教育系统的弊端,在生命最后的二十年里他把所有的精力都投入到数学教育的改革中,特别是对中学教育.他还强调应该鼓励年轻学生利用他们的直觉去解决问题,而不是仅仅学些与他们的生活毫无联系的技巧和结论.

习　题

集合论中的问题

1. 下面是理查德悖论(按其提出人 Jules Richard(1862—1956) 命名):将所有两个字母的组合,按字母次序排列,然后是所有三个字母的组合,等等,在此之后消去所有没有界定一个实数的组合(例如,"six" 界定了一个实数而"sx"则没有).那么,以有限个字母界定了实数的集合构成一个可数的良序集 $E = \{p_1, p_2, \cdots\}$.现在定义0和1之间的实数 $s = 0.a_1a_2\cdots$,要求当 p_n 的第 n 位数不为8或9时,a_n 等于 p_n 的第 n 位数加1,否则就等于1.尽管 s 是由有限个字母界定但却不在 E 中.是个矛盾.如何消解此悖论?此悖论是与理发师悖论有关还是与源于策梅洛的悖论有关?
2. 证明三分法可由策梅洛的良序定理推导出.
3. 证明策梅洛的分离公理消解了罗素的理发师悖论以及理查德悖论(习题1),这里消解的意思是说,某些"集合"现

在被排除在讨论之外了.

点集拓扑中的问题

注:本节中的问题要求对拓扑的基本概念有一定程度的熟悉.

4. 写出并证明平面上的海涅 - 波莱尔定理.
5. 证明在杨夫妇意义下的一个连通集 A 不能表示为 $A = B \cup C$,其中 B 和 C 为闭且 $B \cap C = \varnothing$.
6. 证明[0,1] 中的有理数集是不连通的.
7. 设集合 E 按弗雷歇用集合套定义为紧.证明 E 的每个无限子集 E_1 至少有一个在 E 中的极限点.
8. 用例子证明实直线上的集合套性质由这些子集为既闭又有界而定.
9. 证明一个定义在闭且紧的集合 E(依弗雷歇的定义)上的实值函数为序列式连续时,则它在其上有界并至少一次达到它的上界.
10. 证明:如果 E 按弗雷歇的定义为闭且紧,而且如果$\{E_n\}$是 E 的闭子集构成的集合套,则交集 $\bigcap_n E_n$ 非空.
11. 证明$[a,b]$上的实连续函数空间在极大模度量下为弗雷歇意义下的"正规".
12. 证明:所有实数无穷序列$\{x = \{x_1, x_2, \cdots\}\}$的空间在度量:
$$(x,y) = \sum_{p=1}^{\infty} \frac{1}{p!} \frac{|x_p - y_p|}{1 + |x_p - y_p|}$$
下为正规.
13. 证明:在练习 12 中定义的度量空间中存在一个数 a,使得对此空间中所有的点 x, y,满足$(x,y) < \alpha$.
14. 证明:如果 E 为拓扑空间,A 为其闭子集,即包含了它的所有聚点的子集,则 $E - A$ 是按豪斯多夫概念下的区域(即开集).反之,如果 E 的子集 B 为开,用豪斯多夫的定义证明 $E - B$ 为闭.
15. 证明:在豪斯多夫的极限点的定义下,一个给定的无限集合 A 可以没有一个极限点.
16. 证明:豪斯多夫的两个关于在一点连续的定义是等价的.
17. 运用豪斯多夫关于连续性的邻域定义证明:一个连续函数保持了连通性和紧性.

组合拓扑中的问题

18. 在球面上有多少个线性无关的闭一维子簇?在环面上呢?
19. 确定图 18.5 中画出的四面体 $V_0 V_1 V_2 V_3$ 的边缘,证明此边缘的边缘为 0.
20. 对上题中四面体的面 $V_0 V_1 V_3$ 计算其边缘,并证明此边缘的边缘为 0.

代数中的问题

21. 运用迪克森对域的公理推导出下列定理:
    (a) 对此集合中任意两个元 a, b,则在此集中存在元 y,使得 $a + y = b$.
    (b) 如果对某个特定元 a 有 $a + z = a$,则对每个元 b 成立 $b + z = b$.
    (c) 如果 $a + b = a + b'$,则 $b = b'$.
22. 证明:除去迪克森公理中第一个外,集合$\{0,1,-1\}$在通常的加法和乘法下满足其余的各条公理.
23. 证明:集合 $s = \{r\sqrt{2} \mid r$ 为有理数$\}$在通常加法和乘法下,除去第 5 条外满足迪克森公理中其余各条.在此情形下,找出满足$(a \times x) \times b = b$ 的x.
24. 证明:在相对于 $p = 5$ 的亨泽尔的乘法中,乘积 4.324 × 3.403 等于 2.2312242.
25. 证明:在相对于 $p = 5$ 的亨泽尔的乘法中,如果进行长除的话,则 3.12 ÷ 4.21 的商数为循环数 2.42204220….
26. 证明:在亨泽尔的 5 进制数域中,$3.12 \div 0.2 = 4 \cdot 5^{-1} + 0 + 1 \cdot 5$.
27. 证明:对于亨泽尔的 5 进制数;一个单位(即其最低 p 的幂为零的非零数)的乘法逆元仍为单位.
28. 设 x 为一个 p 进制数,对一个整数 r 定义 x 的一个邻域 $U_r(x)$ 为 $U_r(x) = \{y \mid y = x(\bmod p^r)\}$.证明对这样选

取的 x 的邻域使域 $\mathbf{Q}_p$ 成为在豪斯多夫意义下的拓扑空间.

29. $\mathbf{Q}_p$ 中任一个数 x 可惟一地写作 $x = p^{\alpha}e$,其中 e 是单位.称整数 α 称作 x 相对于 p 的阶,记作 $v_p(x)$.对 $\mathbf{Q}_p$ 中任意的 x, y,定义 (x, y) 为 $(1/p)^{v_p(x-y)}$.证明 (x, y) 在 $\mathbf{Q}_p$ 上定义了一个弗雷歇意义下的度量.
30. 运用习题 29 中的度量定义 $\mathbf{Q}_p$ 中通常的柯西序列的观念,证明 $\mathbf{Q}_p$ 中每个柯西序列收敛于一个极限.
31. 证明下面两个基元的乘法表中的每个都定义了实数上的一个二阶结合代数.还有其他的这种乘法表吗?

	i	j
i	i	j
j	j	0

	i	j
i	i	j
j	0	0

	i	j
i	j	0
j	0	0

32. 举出不同于正文中的范畴的多个例子.利用你所列出的那些范畴以及正文中范畴给出许多函数的例子.

计算机方面的问题

33. 证明一个 n 次多项式的 n 次差分总为常数.
34. 制作一张差分表,它使巴贝奇的差分机能够计算棱锥数,即三角形数的和的那些数,它们可以想成譬如是在已知高的三棱锥中能容纳的加农炮炮弹的数目.
35. 证明 14 章习题 3 中定义伯努利数的方程可以转换为阿达·洛夫莱斯用过的方程,她用此方程写出了计算这些数的程序.(提示:利用 e^x 的幂级数.)
36. 考虑一部具两个状态 q_1, q_2 且能够打印出 0 和 1 符号的图灵机.假定它由下述指令定义:

 (a) 如果机器处于状态 q_1 并读出 1,则它打印 1,向右移动一格并保持在状态 q_1.

 (b) 如果机器处于状态 q_1 并读出空格,则它打印 1,向右移动一格,并变到状态 q_2.

 (注意,当它处于状态 q_2 时对机器没有任何指令.)假设机器开始状态为 q_1 且纸带的第一格为空白,而其右边的下两格都为 1,而其余的右端方格均为空白,另外更进一步假定最左端的 1 就是读出的那个最初的方格.证明纸带的最终构形与最初的那个相同,但当它有三个 1 而不是两个时则不相同.一般来说,一个有 n 个 1 的纸带被解释为代表了数 $n + 1$.证明这样的图灵机可对任意非负整数 n 计算函数 $f(n) = n + 1$.
37. 确定一个计算函数 $f(n) = 2n$ 的图灵机.
38. 证明下面的布尔展开定理:$f(x, y) = [f(0, y) + x][f(1, y) + x']$,其中 f 为两个布尔变量 x, y 的布尔函数.
39. 运用加法对乘法的分配律 $a + (bc) = (a + b)(a + c)$ 及习题 38 中的定理建立下面的布尔函数的法则:

$$x + f(x, y) = x + f(0, y).$$

40. 证明布尔展开定理 $f(x, y) = xf(1, y) + x'f(0, y)$,以及定理 $xf(x, y) = xf(1, y)$.注意一下,这些是和定理 38 和 39 中结果的对偶.
41. 像正文中描述的那样,构造一个代表二进位数加法的电路.

讨论题

42. 查验一下正文中关于集合论的巴拿赫 - 塔尔斯基悖论并讨论其意义和所暗含的内容.你相信此结果吗?这与你对选择公理真实性的信任是如何关联的?
43. 概述在现代代数课上的阐述亨泽尔 p 进数的一系列课程,指出在这样的域上如何既能展开代数学又能展开分析学.
44. 是否能够在今天的拓扑课中使用豪斯多夫的 1914 年的课本.找出一本来并将其与当今的课本作出比较.
45. 找一本范德瓦尔登的《近世代数》的早期版本.将它与当今的课本作比较.
46. 读一两篇谈及阿佩尔 - 哈肯四色定理证明的文章.讨论这个证明的含意.你相信这个推理证明了定理吗?

文献和注解

虽然没有易于使用的 20 世纪数学通史,然而本章中每个专题都有一本以上的好书论述过. Gregory H. Moore 的 *Zermelo's Axiom of Choice: Its Origins, Development and Influence*(New York: Springer, 1982) 论述了本章第一部分的材料. 参考的许多文章的原始文本可在 Jean van Heijenoort 编纂的 *From Frege to Gödel: A Source Book in Mathematical Logic*, 1879—1931(Cambridge: Harvard University Press, 1967) 中找到. 点集拓扑学的一部通史是 Jerome Manheim 的 *The Genesis of Point Set Topology*(New York: Macmillan, 1964). 代数拓扑的历史在 Jean Dieudonné 的 *A History of Algebraic and Differential Topology*, 1900—1960(Boston: Birkhäusr, 1989) 中有十分详尽的讨论. Raymond Wilder 的文章"Evolution of the Topological Concept of 'Connected'," *American Mathematical Monthly* 85(1978), 720—726 和 J.-P. Pier 的文章"Historique de la notion de compacité," *Historia Mathematica* 7(1980), 425 - 443 讨论了点集拓扑中两个重要概念的历史. B. L. Van der Waerden, *A History of Algebra from al-Khwarizmi to Emmy Noether*(New York: Springer, 1985) 的第三部分讲述了 20 世纪代数的许多部分,特别包括了代数论. 在论述计算的历史的方方面面的书中有 B. V. Bowden 编纂的 *Faster Than Thought*(*London*: *Pitman*, 1953),它包括胡洛夫莱斯女士的一部翻译本及注释文,还有在战后那几年英国计算机方面的资料. Herman Galdstine 的 *The Computer from Pascal to von Neumann*(Princeton: Princeton University Press, 1972) 主要讲述了作者在 1940 年代和 50 年代所参与的工作,而 N. Metropolis 等人编纂的 *A History of Computing in the Twentieth Century*(New York: Academic Press, 1980) 包括了 1976 年会议的条刊,此会上许多计算方面的先锋人物作了演讲.

1. Auguste Dick, *Emmy Noether*, 1882—1933(Boston: Birkhauser, 1981), p. 130.
2. Felix E. Browder 编纂 *Mathematical Developments arising from Hilbert's Problems*(Providence: American Mathematical Society, 1976), p. 9. 此著作包括有 Hilbert 的演说以及对从 1900 年到作者行文前这段时间中 Hilbert 每个问题的进展的讨论文章.
3. 引自 B. Rarg 和 W. Thomas 的"Zermelo's Discovery of the 'Russell Paradox'," *Historia Mathematica* 8(1981), 15 - 22, pp. 16 - 17.
4. Zermelo, "Beweis, dass jede Menge Wohlgeordnet Werden kann," *Mathematische Annalen*, 59(1904), 514 - 516,并被翻译、收录在 Van Heijenoort 的 *From Frege to Gödel*, 139 - 141, pp. 139 - 140.
5. 同上, p. 141.
6. Zermelo, "Untersuchungen über die Grundlagen der Mengenlehre I," *Mathematische Annalen* 65(1908), 261 - 281,译文在 Von Heijenoort, *From Frege to Gödel*, 199 - 215, pp. 201 - 204.
7. 引自 Moore 的 *Zermelo's Axiom of Choice*, p. 263.
8. Moore 的 *Zermelo's Axioom of Choice* 有对此观点很好的讨论.
9. Young 和 Young, *The Theory of Sets of Points*(New York: Chelsea, 1972), p. 204. 这是对原始文章的重印.
10. 引自 Michael Bernkopf 的"The Development of Function Spaces with Particular Reference to Their Origisn in Integral Equation Theory",《*Archive for History of Exact sciences*》, 3(1966/67), 1 - 96, p. 37.
11. Fréchet, "Generalisation d'un théoreme de Weierstrass," *Comptes Rendus*, 139(1904), 848 - 850.
12. Hausdorff, *Grundzüge der Mengenlehre*(New York: Chelsea, 1949), p. 213. 这个 Chelsea 的重印本给出了原著的完整内容.
13. 同上, p. 232.
14. 同上, p. 244.
15. 同上, p. 359.
16. Steinitz, "Algebraische Theorie der körper," *Jownal für die Reine und Angewandte Mathematik*, 137(1910), 167 - 310, p. 167.
17. Hermann Wegl, "*Space-Time-Matter*", Henry Brose 翻译. (New York: Dover Publications, 1952), p. 19.

18. Stefan Banach, "Sur les opérations dans les ensembles abstraltes et leur application aux équations intégrales", *Fundamenta Mathematicae*, 3(1922), 133 – 181, p.134. 这篇被引用在 Gregory H. Moore 的"The Axiomatization of Linear Algebra: 1875—1940", *Historia Mathematica*, 22(1995), 262 – 303 中,后文包含有关于向量空间的抽象概念如何发展的更多细节.

19. 引自 Constance Reid 的 *Hilbert*(New York: Springer-verlag, 1970), p.143. 关于诺特的数学及它引起的发展的更多信息可在 James W. Brewer 和 Martha K. Smith 编纂的 *Emmy Noether: Atribute to her life and work*(New York: Marcel Dekker, 1981) 中找到.

20. Dick, *Emmy Noether*, pp, 173 – 174. 在此卷中可找到亚里克山大罗夫的完整读话内容以及范德瓦尔登的纪念读话.

21. Eilenberg 和 Mac Lane, "General theory of nature equivalence," *Transaction of the American Mathematical Society*, 58(1945), 231 – 294), p.237.

22. 参阅 Morris Kline 的 *Mathematics: The Loss of Certainty*(New York: Oxford University Press, 1980), 有关于许多近代数学内容过于抽象方面的激烈争辩. 另一方面, 关于甚至抽象数学的发展也有实用性的观点可参见 Eugene P. Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences," *Communications in Pure and Applied Mathematics* 13(1960), 1 – 14.

23. Leibniz, 在 David Eugene 的 *A Source Book in Mathematics*(New York: Dover, 1959), pp.180 – 181.

24. 参见 Allan G. Bromley, "Charles Babbage's Analytical Engine, 1838" *Annals of the History of Computing* 4(1982), 196 – 217, 有更多的细节.

25. Ada Lovelace, 她翻译的 L.F. Menabrea 的文章"Sketch of the Analytical Engine Invented by Charles Babbage" 中的, "Note G" 此文收录在由 Philip 和 Emily Morrison, eds., *Charles Babbage: On the Principles and Development of the Calculator and Other Seminal Writing by Charles Babbage and Others*(New York; Dover, 1961), 225 – 297, p.284.

26. 引自 Goldstine 的 *The Computer*, pp, 243 – 244.

27. 纠错码学科的历史以及其因对它们的研究而引起的数学的发展可参见 Thomas Thompson, *From Error Correcting Codes Through Sphere Packings to Simple Groups*(Washington: Mathematical Assocoation of America, 1983).

28. 参见 George B. Dantzig *Linear Programming and Extensions*(Princeton: Princeton Untuersity Press, 1963), 它有关于线性规划及其历史进程的更多信息. 它也包括了在此领域中原始文章的大量参考资料.

29. 此信件引自 Norman Biggs, E. Keith Lloyd, 和 Robin J. Wilson 的 *Graph Theory*, 1736—1936(Oxford: Clarendon, 1976), pp.90 – 91. 此书还包括了许多图论方面的原始文章, 其中有关于四色问题的. 对此问题更新近的历史可见 Thomas L. Saaty 和 Paul C. Kainen 的 *The Four-Color Problem: Assaults and Conquest*(New York: Dover, 1986). 对计算机辅助的四色问题证明的哲学含意的讨论可参见 Thomas Tymoczko 的两篇文章: "Computers, Proofs and Mathematicians: A Philosophical Investigation of the Four-Color Proof," *Mathematics Magazine* 53(1980), 131 – 138 和 "The Four-Color Problem and its Philosophical Significance," *Journal of Philosophy* 76(1979), 57 – 83.

20世纪数学概览

1792—1871	巴贝奇(Charles Babbage)	分析机
1809—1880	皮尔斯(Benjamin Peirce)	线性结合代数
1815—1852	艾达·金(Ada Byron King)	计算机程序
1821—1895	凯莱(Arthur Cayley)	树
1845—1918	康托尔(Georg Cantor)	良序原理
1854—1912	庞加莱(Henri Poincaré)	代数拓扑
1861—1941	亨泽尔(Kurt Hensel)	p 进位数
1863—1942	威廉·杨(William Young)	集合论的教科书
1868—1942	豪斯多夫(Felix Hausdorff)	豪斯多夫空间
1868—1944	格瑞斯·杨(Grace Chisholm Young)	集合论的教科书
1871—1928	斯泰尼茨(Ernst Steinitz)	域论
1871—1953	策梅洛(Ernst Zermelo)	集合论公理
1871—1956	埃米·波莱尔(Emile Borel)	海涅 - 波莱尔定理
1872—1970	罗素(Bertrand Russell)	罗素悖论
1874—1954	迪克森(Leonard Eugene Dickson)	域的公理
1878—1973	弗雷歇(Maurice Fréchet)	拓扑的基本概念
1882—1935	埃米·诺特(Emmy Noether)	环论
1882—1948	韦德伯恩(Joseph H. M. Wedderburn)	任意域上的代数
1885—1955	外尔(Hermann Weyl)	向量空间的公理
1887—1948	梅尔(Walther Mayer)	同调论的公理
1888—1971	亚历山大(James Alexander)	单纯形
1891—1965	弗伦克尔(Abraham Fraenkel)	集合论的公理
1892—1945	巴拿赫(Stefan Banach)	巴拿赫 - 塔尔斯基悖论、向量空间
1894—1971	霍普夫(Heinz Hopf)	同调群
1896—1982	亚历山大罗夫(Pavel Sergeiivich Aleksndrov)	代数拓扑
1901—1983	塔尔斯基(Alfred Tarski)	巴拿赫 - 塔尔斯基悖论
1903—1957	冯·诺依曼(John von Neumann)	计算机设计
1906—1978	哥德尔(Kurt Gödel)	相容性的不可能性证明
1906—1993	佐恩(Max Zorn)	佐恩引理
1907—1989	惠特尼(Hassler Whitney)	对偶图
1909—	麦克莱恩(Saunders Mac Lane)	范畴论
1912—1986	康托洛维奇(Leonid V. Kantorovich)	线性规划
1912—1954	图灵(Alan Turing)	图灵机
1913—	艾伦伯格(Samuel Eilenberg)	范畴论
1914—	丹奇克(George Dantzig)	线性规划
1915—	海明(Richard Hamming)	纠错码
1916—	香农(Claude Shannon)	开关电路
1928—	哈肯(Wolfgang Haken)	四色定理
1932—	阿佩尔(Kenneth Appel)	四色定理
1934—	科恩(Paul Cohen)	选择公理的独立性

习题答案

第 1 章

2. |||/|| ∩∩ 9; ΥΥ ΥΥΥ/ΥΥ

3. $125 = \rho\kappa\varepsilon, 62 = \xi\beta, 4821 = {}^{,}\delta\omega\kappa\alpha, 23\,855 = M^{\beta'}\gamma\omega\nu\varepsilon$.

4. [illegible]

5. $16\ \bar{2}\ \overline{420}$(还有以埃及分数表示此答案的其他可能方式).

6. $99\ \bar{2}\ \bar{4}$.

7. $2 \div 11$:

1	11
$\bar{\bar{3}}$	$7\ \bar{3}$
$\bar{3}$	$3\ \bar{\bar{3}}$
$\bar{6}$	$1\ \bar{\bar{3}}\ \bar{6}'$
$\overline{66}$	$\bar{6}'$
$\bar{\bar{6}}\ \overline{\overline{66}}$	2

$2 \div 23$:

1	23
$\bar{\bar{3}}$	$15\ \bar{3}$
$\bar{3}$	$7\ \bar{\bar{3}}$
$\bar{6}$	$3\ \bar{2}\ \bar{3}$
$\overline{12}$	$1\ \bar{2}\ \bar{4}\ \bar{6}'$
$\overline{276}$	$\overline{12}'$
$\overline{\overline{12}}\ \overline{\overline{276}}$	2

8. $5 \div 13 = \bar{4}\ \overline{13}\ \overline{26}\ \overline{52}; 6 \div 13 = \bar{4}\ \overline{26}\ \overline{52} = \bar{4}\ \bar{8}\ \overline{13}\ \overline{104}$.

11. $18 \leftrightarrow 3,20; 32 \leftrightarrow 1,52,30; 54 \leftrightarrow 1,06,40; 1,04 \leftrightarrow 56,15$.

12. $25 \times 1,04 = 26,40; 18 \times 1,21 = 24,18; 50 \div 18 = 50 \times 0;03,20 = 2;46,40; 1,21 \div 32 = 1,21 \times 0;01,52,30 = 2;31,52,30$.

13. $16\ \bar{2}\ \bar{8} = 16\frac{5}{8}$.

15. $51\frac{41}{109}, 32\frac{12}{109}, 16\frac{56}{109}$.

16. $\frac{9}{25}, \frac{7}{25}, \frac{4}{25}$.

17. 共有 7 人,价格为 53.

18. 一天的$\frac{15}{74}$.

19. 10.9375 磅.

20. $90°$ 角:$\frac{1}{4}, \frac{\pi}{2} - \frac{1}{2}$;$60°$ 角:$\frac{11 - 6\sqrt{3}}{8} = .076, \frac{\pi}{6} - \frac{\sqrt{3}}{4} = 0.091$;$45°$ 角:.032,.039.

24. 第一种情形下,正确的公式给出 $V = 56$,而巴比伦公式给出 $V = 60$,误差 7%.第二种情形下,正确公式给出 $V = \frac{488}{3} = 162\frac{2}{3}$,而巴比伦公式给出 $V = 164$,误差 0.8%.

29. $1;24,51,10 = 1.414212963; \sqrt{2} = 1.414213562$.

30. $1 \div 1;45 \approx 0;34,17,09; \sqrt{3} \approx 1;43,55,42$.

31. $12\ \bar{\bar{3}}\ \overline{15}\ \overline{24}\ \overline{32} = 12\frac{129}{160}; \left(12\frac{129}{160}\right)^2 = 164.0000391$.

33. 直线 6:$v + u = 2\frac{2}{9} = 2;13,20$;直线 13:$v + u = 1\frac{7}{8} = 1;52,30$.

34. (67319,72000,98569).

35. 50.5.

37. 30,25.

38. 8,6.

39. 正方形的边为 250 *pu*.

42. $x = 3\frac{1}{2}, y = 2\frac{1}{3}$.

第2章

4. $n^2 = \frac{(n-1)n}{2} + \frac{n(n+1)}{2}$.　　5. $8 \cdot \frac{n(n+1)}{2} + 1 = 4n^2 + 4n + 1 = (2n+1)^2$.

9. 例子:(a)(3,4,5),(5,12,13),(7,24,25),(9,40,41),(11,60,61);
(b)(8,15,17),(12,35,37),(16,63,65),(20,99,101),(24,143,145).

13. $4ax + (a-x)^2 = (a+x)^2$.　　22. 9;1.

24. 对46;6,计算为$46 = 7 \cdot 6 + 4$;$6 = 1 \cdot 4 + 2$;$4 = 2 \cdot 2$.对23;3,计算为$23 = 7 \cdot 3 + 2$;$3 = 1 \cdot 2 + 1$;$2 = 2 \cdot 1$.

33. ab 为 a^2 与 b^2 之间的比例中项.　　34. a^2b 和 ab^2 为 a^3 与 b^3 两个比例中项.

41. 圆周长 = 250 000斯塔德 = 129 175 000英尺 = 24 456公里;直径 = 79577.5斯塔德 = 41 117 680英尺 = 7787英里.

第3章

1. 离较重的一边 $4\frac{1}{6}$ 米.　　2. 斜向较重的一边.

6. 经十次叠代,在我的计算器上生成的 π 值为3.141593746. π 的到9位的实际值为3.141592654.你的计算器可能会生成稍微不一样的结果.

18. $x^2 = 4ay, y(3a - x) = ab$.

21. $y^2 = px$ 的焦点在$\left(\frac{p}{4}, 0\right)$,因此正焦弦长为$2\sqrt{p \cdot \frac{p}{4}} = 2 \cdot \frac{p}{2} = p$.

22. 此椭圆方程可写为$\frac{(x-a)^2}{a^2} + \frac{y^2}{pa/2} = 1$,因此 $b^2 = pa/2$.

第4章

1. crd 30° = 31;3,30;crd 15° = 15;39,47;crd $7\frac{1}{2}$° = 7;50,54;crd 120° = 103;55,23;crd 150° = 115;54,40;
crd 165° = 118;58,25;crd $172\frac{1}{2}$° = 119;42,28.

4. crd 12° = 12;32,36.　　9. 在纬度40°,影长为50;21.在纬度$23\frac{1}{2}$°,影长为26;5.

11. 夏至:在纬度36°,影长为13;18.在纬度$23\frac{1}{2}$°,影长为0.
冬至:在纬度36°,影长为101;52.在$23\frac{1}{2}$°,影长为64;21.

12. 45° : δ = 16°37′;α = 42°27′;　315° : δ = − 16°37′;α = − 42°27′;
90° : δ = 23°51′;α = 90°;　270° : δ = − 23°51′;α = − 90°;
120° : δ = 20°30′;α = 122°16′;　240° : δ = − 20°30′;α = − 122°16′.

14. λ = 60°,ρ = 35°47′;λ = 90°,ρ = 63°45′.

15. 在纬度36°白天的长为211°32′ = 14小时零6分,因此日出在4:57 A.M,日落在7:03 P.M.在纬度45°,白天的长为223°54′ = 14小时56分,因此日出在4:32 AM.日落在7:28 P.M.

16. 纬度是40°53′,夏至日出的位置是东北32°20′,日落的位置是西北32°20′.在冬至日出的位置是东南32°20′,日落的位置是西南32°20′.

18. λ = 45°:太阳离天顶为28°23′;λ = 90°:太阳离天顶21°9′.

19. 大约为5月20日和7月21日.

20. 在纬度45°,北向日落点为东北34°53′.在纬度20°为东北25°29′.子夜太阳大约开始于纬度75°的4月30日.

22. 10.9.　　23. $\sqrt{3} \approx \frac{26}{15}$.　　25. $\sqrt{23} \approx \frac{43}{9}$.

第5章

1. 第 n 个五边形数为$\frac{3n^2-n}{2}$,第 n 个六边形数为 $2n^2-n$.

2. 第 n 个三角锥数为$\frac{n(n+1)(n+2)}{6}$,第 n 个正方形放的锥形数为$\frac{n(n+1)(2n+1)}{8}$.

7. 84.　8. 12,8.　9. $72\frac{1}{4}$, $132\frac{1}{4}$.　11. $\frac{121}{16}$.

12. 13,3.　13. 12,8.　15. $x=\frac{5}{7}$, $y=\frac{267}{343}$.　16. $\frac{2481}{7921}$, $\frac{5440}{7921}$.

22. $2\pi^2r^2R$.　23. 336.　24. $\frac{12}{25}$天.　25. $A:15\frac{5}{7}$, $B:18\frac{4}{7}$.

第6章

2. 57.5.　3. 9.　4. 24600.

5. 一个解是:1 品脱高质量的葡萄酒和 9 品脱的酒渣.

7. (a)6.35; (b)20.

8. 36.　9. 60.　10. 12.　11. 24.　13. 23.

14. $d=12$, $C=36$.　15. sin 15° = 890; sin 18°45′ = 1105; sin 22°30′ = 1315.

17. 237.　18. $x=2$, $y=1000$, $N=3000$.

19. $x=41$, $y=94$, $N=5640$.　20. $x=2$, $y=731$; $x=20$, $y=7310$.

22. 59.　24. $m=12$, $n=53$.　26. $x=9$, $y=82$.　28. $x=180$, $y=649$.

31. 148,608.　32. $\frac{1}{14}$天; $\frac{2}{14}$, $\frac{3}{14}$, $\frac{4}{14}$, $\frac{5}{14}$.　33. $8\frac{3}{4}$.

34. 一个解是:9 只孔雀,21 只鸽子,7 只天鹅,30 只萨拉沙鸟.

35. 一个解是:第一个旅客有 11 个钱币,第二个有 13 个钱币,而钱包有 30 个钱币.

第7章

4. (a)12;(b)3;(c)24.　5. (a)4;(b)3.　6. 6,4.

7. (a)$x=\sqrt{2\frac{1}{2}+\sqrt{1000}}-\sqrt{2\frac{1}{2}}$, $y=10+\sqrt{2\frac{1}{2}}-\sqrt{2\frac{1}{2}+\sqrt{1000}}$;

　(b)$x=15-\sqrt{125}$, $y=\sqrt{125}-5$.

8. (a)$x=\left(1+\sqrt{2}+\sqrt{13+\sqrt{8}}\right)^2$;(b)$x=4\frac{1}{2}-\sqrt{8}$;(c)$x=3,16\frac{1}{3}$.

9. $x=\sqrt[4]{\sqrt{12500}-50}$, $y=10/x$, $z=100/x^3$.

15. 例子:$c=2$, $d=2$ 不相交;$c=3$, $d=2$ 给出一个交点;$c=4$, $d=2$ 给出两个交点.

21. $17\,296=2^4p_3p_4$, $18\,416=2^4q_4$.

29. 124°32′.　31. 13 331 731 肘尺,或者近似于 3787 英里.

32. 如果 $AB=60°$, $AC=75°$, $BC=31°$,则 $\angle A=29°32'$, $\angle B=112°25'$ 和 $\angle C=55°59'$.

第8章

1. 3600,2400,1200.　3. 375 步.　4. 6;23,21 = 6.389.　5. 3.848.　6. 5.47.

7. 10.　8. $l=8$, $w=6$, $A=48$.　9. 10.　10. 边长 1 的正五边形的面积为 1.72.　11. 50.

12. 10.　13. $12\frac{1}{2}$.　25. $1\frac{23}{37}$小时.　27. 119.

28. $25+10\sqrt{5}, 20+10\sqrt{5}$. **30.** $\left(\frac{25}{24}\right)^2$. **31.** $x=5, y=4$ 或者 $x=6, y=3$.

32. $x=6, y=4$. **33.** $x=6, y=3$.

插入章

2. (11,20,88).

3. 设 $\Delta t = t_1 - t_0 \pmod{13}$, $\Delta v = v_1 - v_0 \pmod{20}$ 和 $\Delta y = y_1 - y_0 \pmod{365}$,则最小的天数为 $365[40\Delta t - 39\Delta v - \Delta y] + \Delta y \pmod{18980}$.

4. 1,8,15,18 + 2,12,13,0 = 4,1,10,18 = 29,378 天. Pacal 活到他的第 81 岁.

5. e 部分的一位妇女在 m 部分有个母亲,在 f 部分有个父亲,在 mf 部分有位丈夫而在 m^2 部分是其孩子们.

第 9 章

1. 一个金币的$\frac{11}{90}$. **2.** $133\frac{3}{8}$. **3.** 32.6 镑.

4. 他们将在 $3\frac{15}{16}$ 天相遇;从罗马来的信使走了 $140\frac{5}{8}$ 走了 $140\frac{5}{8}$ 里,而从威尼斯来的走了 $109\frac{3}{8}$.

5. 第三个合股人投资 $103\frac{8}{93}$;第二个合股人获利 129,第三个合股人获利 153.

6. 第一人:10 天;第二人:24 天;第三人:$17\frac{1}{7}$ 天.

7. $3\frac{1}{3}$ 小时. **8.** $\frac{\sqrt{43}}{2}+\frac{\sqrt{11}}{2}, \frac{\sqrt{43}}{2}-\frac{\sqrt{11}}{2}$. **9.** $5\frac{359}{389}, 4\frac{30}{389}$.

10. 大约每月每里拉 2.5 迪拉里尼斯. **15.** $\frac{2}{7}\sqrt[3]{1225}, \frac{2}{5}\sqrt[3]{1225}$. **17.** $\frac{4}{3}$ 小时.

18. 女儿:$14\frac{2}{7}$;母亲:$28\frac{4}{7}$;儿子:$57\frac{1}{7}$. **19.** $5+\sqrt{2}$. **20.** 80 天.

21. $5-2\sqrt{3-2\sqrt{2}}, 5+2\sqrt{3-2\sqrt{2}}$. **24.** 15 位公爵,450 位伯爵,27 000 名士兵.

25. 3,5. **27.** $\frac{3\pm\sqrt{5}}{2}, -3$. **29.** $x=\sqrt[3]{\sqrt{26}+5}-\sqrt[3]{\sqrt{26}-5}$.

30. $x=\sqrt[3]{4}+\sqrt[3]{2}$. **32.** $5, 2\pm\sqrt{3}$. **33.** $b=3$;两个解为 $x=1\pm\sqrt{3}$.

34. 弗朗西斯:-48;嫁妆:52.

36. $4\pm\frac{4}{3}\sqrt{3}$. **38.** $4+\sqrt{-1}$. **41.** 8,2. **42.** $x=12$.

43. 13⓪3①9②5③;22⓪8①6②4③2④. **44.** 177⓪8①3②9③.

第 10 章

5. 从赤道到 10° 的距离:10.05 cm;从 10° 到 20° 的距离:10.36 cm;从 20° 到 30° 的距离:11.05 cm.

7. $AB = 8.46, AG = 14.10$.

8. $\angle A = 90°; \angle B = 63°; \angle C = 27°; a:b:c = 1.12:1:0.51$.

9. $BC = 15.78; AB = 33.06; AC = 30.06$.

10. 121.41. **11.** 顶角 35°,底角 72.5°. **14.** 1.88 年 = 687 天.

第 11 章

2. 一个椭圆. **4.** $x^2+y^2=\frac{m}{2}-a^2$(m 为所给面积).

6. $(e-cg)y^2+(de+fgc-bcg)xy+bcfgx^2+(dek-fglc)y-bcfglx=0$.

9. $2+\sqrt{7},2-\sqrt{7},-2+\sqrt{2},-2-\sqrt{2}$.　　10. $\frac{2\sqrt{3}}{9},\frac{\sqrt{3}}{3},\frac{4\sqrt{3}}{9}$.

14. $\sin\alpha=\frac{b}{a},\cos\alpha=\frac{\sqrt{a^2-b^2}}{a}$.　　16. $x=-9\pm\sqrt{57}$.　　17. $x=1,4$.

19. $1,16,-4\frac{1}{2}+\frac{1}{2}\sqrt{17},-4\frac{1}{2}-\frac{1}{2}\sqrt{17}$.

24. 四次投掷没有6点的概率为$\left(\frac{5}{6}\right)^4=\frac{625}{1296}$.于是胜率为$(1296-625):625=671:625$.

26. $42:22$.　　27. $\frac{17}{27},\frac{5}{27},\frac{5}{27}$.　　30. 第一个:$\frac{37}{72}$,第二个:$\frac{35}{72}$.

31. $31:30$.　　32. $9:6:4$.　　33. $A:1000$ vs. $B:8139$.

第12章

1. $S=\frac{20\sqrt{3}}{3},V=\frac{8000\sqrt{3}}{9}\approx1539.6$.　　4. $\frac{2b}{3}\sqrt{\frac{b}{3}}$.

7. $t=\frac{pxy-3y^3}{3x^2-py}$.　　9. $t=-\frac{a^2y^2}{b^2x}$.　　10. $v=\frac{3}{2}x_0^2+x_0$.

29. $\frac{1024}{625}\left\{\left[\sqrt{a}\left(\frac{a^2}{5}-\frac{a}{3}\right)\right]+\frac{2}{15}\right\}$,其中 $a=1+\frac{25}{16}\sqrt{b}$.

33. $\sqrt{1+x}=1+\frac{x}{2}-\frac{x^2}{8}+\frac{x^3}{16}-\frac{5x^4}{128}+\cdots$.　　41. $\frac{\dot{y}}{\dot{x}}=\frac{a^2-2x^2}{2y\sqrt{a^2-x^2}}$.

42. $y=4x-2x^2+\frac{x^3}{3}-\frac{x^4}{2}-\frac{2x^5}{5}-\cdots$.

第13章

3. $xy=k$.　　6. $y=x^2+\frac{k}{x}$.　　7. $y=\frac{kx^2}{a-kx}$.

10. $y=Ae^x+Be^{2x}+Ce^{3x}$.　　14. $x^2y^3+2x^3y^2+4x^2+3y=k$.

15. 悬链线:$y=c\cosh\frac{x}{c}$.　　16. 底边 $=2\sqrt{3}$,高 $=3$,边 $=2\sqrt{3}$.

17. $r=\sqrt[3]{\frac{3\sqrt{2}V}{2\pi}},h=\sqrt{2}r$.　　20. 当 $y=\frac{a\sqrt{2}}{2}$.

21. 如果此矩形的长为 a,宽为 b,且如果此直线通过的那个顶点的坐标为(0,0),则所要求的直线通过$(-\sqrt[3]{ab^2},b)$和$(a,-\sqrt[3]{a^2b})$.

23. 近似于$\frac{1}{16}$.　　27. $\left(1,\frac{3}{2}\right)$是个极大点,而$\left(\frac{1}{2},\frac{3}{4}\right)$既非极大也非极小.

36. $\frac{8}{21}a^{3/4}b^{1/2}x_0^{7/4}$.　　40. $p=\frac{1}{2\sqrt{x}},q=-\frac{1}{8x\sqrt{x}},r=\frac{1}{32x^2\sqrt{x}}$.

第14章

1. $B_8=-\frac{1}{30},B_{10}=\frac{5}{66},B_{12}=-\frac{691}{2730}$.

2. $\sum j^{10}=\frac{1}{11}n^{11}+\frac{1}{2}n^{10}+\frac{5}{6}n^9-n^7+n^5-\frac{1}{2}n^3+\frac{5}{66}n$.

8. $\frac{24864}{59049}=0.42$.

9. $x=6.6$,而近似值为6.3.对7次试验,其可能性较均等情形好,而6次试验,它们较均等情形差.

10. 150.　**11.** 28.　**17.** 0.91.　**19.** $\frac{9}{10}$.　**21.** 10.35.

23. $\frac{8}{9}$ 周.　**24.** 54 头牛.　**25.** $\frac{a^2-4b^2}{2a}$.　**26.** 离爱丁堡 200 哩.　**28.** 7.

30. 5 个男人,15 个女人.　**31.** 70 克朗.　**32.** 老三:64;老二:72;老大:84.

33. (1,5,8,12),(2,10,3,11),(4,7,6,9).　**34.** 模 13 的二次剩余为 1,3,4,9,10,12.

38. $3, 2\omega+\omega^2, 2\omega^2+\omega$.　**39.** x_0+y_0.　**40.** 次切线 $=\frac{x^2-2a^2}{x}\sqrt{\frac{2x^2-a^2}{x^2-a^2}}$.

41. $\frac{abx\sqrt{by+4x+a}}{3y\sqrt{4x+a}}$.　**48.** 19 头公牛,1 头母牛,80 只羊.

49. $\frac{28}{5},\frac{68}{5},\frac{12}{5},\frac{192}{5}$.　**50.** 102.65 尺.

第 15 章

2. $2^3\equiv1, 3^6\equiv1, 4^3\equiv1, 5^6\equiv1, 6^2\equiv1$.　**3.** 2,6,7,11.　**8.** $3+5i=(1-4i)(-1+i)$.

15. 如果此群为$\{1,\alpha,\alpha^2,\cdots,\alpha^{17}\}$,则 6 阶循环子群为$\{1,\alpha^3,\alpha^6,\alpha^9,\alpha^{12},\alpha^{15}\}$,而其他的倍集为$\{\alpha,\alpha^4,\alpha^7,\alpha^{10},\alpha^{13}\alpha^{16}\}$和$\{\alpha^2,\alpha^5,\alpha^8,\alpha^{11},\alpha^{14},\alpha^{17}\}$.

16. 设 $\rho=\sqrt[3]{\frac{7}{2}(1+3\sqrt{-3})}$ 和 $\sigma=\sqrt[3]{\frac{7}{2}(1-3\sqrt{-3})}$,并设 $\alpha_1=\frac{1}{3}(-1+\rho+\delta)$, $\alpha_2=\frac{1}{3}(-1+\omega^2\rho+\omega\sigma)$, $\alpha_3=\frac{1}{3}(-1+\omega\rho+\omega^2\sigma)$,其中 ω 是 1 的三次复根.于是 6 个解为 $x=\frac{1}{2}\left(\alpha_i\pm\sqrt{\alpha_i^2-4}\right)$, $i=1,2,3$.

21. S_3.　**22.** 其中三个为阿贝尔群.

23. x^2+x+1 为模 5 不可约.因此$\{\alpha_0+\alpha_1\alpha+\alpha_2\alpha^2\}$是个 5^3 阶的域,其中 α 满足 $\alpha^3=-\alpha-1$,而 $0\leqslant\alpha_j<5$, $j=0,1,2$.

28. $\alpha\beta=12-9i+18j+24k, \frac{\alpha}{\beta}=\frac{23}{15}i+\frac{2}{3}j-\frac{4}{3}k$.

31. $x\bar{y}z=0$;没有干净的,反刍的分蹄的牲畜.

40. $x=\frac{2}{\sqrt{5}}u+\frac{1}{\sqrt{5}}v, y=-\frac{1}{\sqrt{5}}u+\frac{2}{\sqrt{5}}v$.

41. 不为零的行列式的最大阶为 2.表达解的一种方式是:$u=-10x-15z, v=18x-y+27z$,其中 x,y,z 取任意值.

42. 秩 $=2$;此解集的一个基为$(-10,18,1,0,0),(0,-1,0,1,0),(-15,27,0,0,1)$.

43. 伴随方程组为 $-10u+18v+x=0, -v+y=0, -15u+27v+z=0$;解集的一个基为$(1,0,10,0,15)$, $(0,1,-18,1,27)$.

第 16 章

14. $\theta\approx0.46$.　**16.** $\frac{2}{3}$.

31. 例子:$P=\left\{\frac{1}{m}-\frac{1}{m(n+1)}\right\}$,其中 m,n 为正整数,$P'=\left\{1,\frac{1}{2},\frac{1}{3},\frac{1}{4},\cdots\right\}$, $P''=\{0\}$.

38. 第一部分:设 C_1,C_2 为连接三度空间中两点 p_1,p_2 的两条路径.设 C 为一闭路径,它从 p_1 到 p_2 沿 C_1 前进,然后沿 C_2 从 p_2 回到 p_1.于是,$\int_{C_1}\sigma\cdot dr-\int_{C_2}\sigma\cdot dr=\int_C\sigma\cdot dr=\iint_A(\nabla\times\sigma)\cdot da=0$.由此得到

$$\int_{C_1}\sigma\cdot dr=\int_{C_2}\sigma\cdot dr.$$

39. $a=1.53, b=-0.87$.

第 17 章

1. $\frac{1}{2}\begin{vmatrix} \mathrm{d}x & \mathrm{d}y \\ \delta x & \delta y \end{vmatrix} = \frac{1}{2}(\mathrm{d}x\delta y - \mathrm{d}y\delta x)$. **3.** $k = \frac{4}{(1+4x^2+4y^2)^2}$.

10. 半径 K 的球面上一个半径 r 的圆的周长为 $2\pi K\sin\frac{r}{K}$.

22. (3,4,1),(-1,7,1). **23.** (1,2,0). **24.** (3,1),(2,-1).

25. $-x+2y+1=0$. **29.** $-18[ijk]$.

第 18 章

18. 1;2.

19. 边缘:$V_1V_2V_3 - V_0V_2V_3 + V_0V_1V_3 - V_0V_1V_2$;边缘的边缘:$V_2V_3 - V_1V_3 + V_1V_2 - V_2V_3 + V_0V_3 - V_0V_2 + V_1V_3 - V_0V_3 + V_0V_1 - V_1V_2 + V_0V_2 - V_0V_1 = 0$.

20. 边缘是 $V_1V_3 - V_0V_3 + V_0V_1$;边缘的边缘是 $V_3 - V_1 - V_3 + V_0 + V_1 - V_0 = 0$.

22. $1+1=2$,它不在此集合中.

23. 如果 $a = r\sqrt{2}$,则 $x = \frac{1}{2r}\sqrt{2}$.

总参考文献

本书的每章都有带有注释的参考文献一节，这些文献对提供那章材料的进一步信息是有用的. 但是，一般来说，如果一个人想要学习数学中一特殊专题的历史，那么从下列著作之一开始他的研究是有好处的.

1. Ivor Grattan-Guinness 编纂，《*Companion Encyclopedia of the History and Philosophy of the Mathematical Sciences*》(London: Routledge, 1994). 这个两卷本的百科全书包括了关于数学的历史和哲学方面的 180 个专题的简短(有时却过于简短) 文章，而每篇都是由这个领域的一个专家写的. 此百科全书在那些通常被认作为应用数学专题上特别擅长，譬如像力学、物理学、工程和社会科学，而在数学史中更为标准的那些专题则相对较弱. 然而，它确是开始数学史中某个专题研究的好的原始资料.

2. Morris Kline，《*Mathematical Thought from Ancient to Modern Times*》(New York: Oxford University Press, 1972). 此书是数学史的新近著作中最全面的一本，而且对 19 和 20 世纪予以了特别关注. 它还给出了文献目录的章节以便提供进一步帮助. 但是它完全缺少了关于中国数学的资料，而关于印度和伊斯兰世界数学的信息也非常简短.

3. Ivor Grattan-Guinness，《*The Fontana History of the Mathematical Sciences*》(London: Fontana Press, 1997). 这本新的一卷本的历史书虽不像 Kline 著作那样巴罗万象，却包括了最新的资料. 作者很好地利用了他所编纂的那本百科全书，并将专家们的著作融合为一个和谐的整体. 像那本百科全书一样，本书也特别着重于 19 世纪的应用专题.

4. Charles C. Gillispie 编纂，《*Dictionary of Scientific Biograph*》(New York: Scribners, 1970—1990). 这本 18 卷的百科全书(包括两个最新的补充卷) 实质上是本按传记形式组织的综合科学史. 它有关于几乎本书提及每个数学家的文章，自然也关于有本书没有提及的许多数学家的文章. 也有一些特别的关于埃及、巴比伦、印度、玛雅的数学和天文学方面的论述文章. 这本书还给出了一个数量众多的索引，以便让人们开始着手一个数学课题，并找到那些考虑过它的所有数学家的文献.

5. Kenneth O. May，《*Bibliography and Research Manual of the History of Mathematics*》(Toronto: University of Toronto Press, 1973). 本书是对从 19 世纪中到大约 1970 年所写的关于数学史方面的说明性和研究性文章的全面目录索引(包括有 30 000 条以上的条目). 它不只是按传记形式组织，还按照了数学专题和历史的分类来组织材料. 虽说它包罗广泛，但参考的文献却没有标题(为了节省篇幅)，故而人们不总能弄明白文章所参照的准确题目.

6. Joseph W. Dauben，《*The History of Mathematics from Antiquity to the Present: A Selective Bibliography*》(New York: Garland, 1985). 此目录索引虽然收录的只有 2000 多条参考文献，远较 May 的书的少. 但却较容易使用，因为它被仔细地加以注解. 另外，只有对一指定科目的"最好" 的工作才被收录(判断是否最好是由编者和他的班子决定的). 无论如何，这很可能是对一个特别课题开始做调研的最好地方.

其他一些可用于参看的数学史的标准著作有 David E. Smith，《*History of Mathematics*》(New York: Dover, 1958), Eric T. Bell，《*The Development of Mathematics*》(New York: M. Graw-Hill, 1945), Eric T. Bell，《*Men of Mathematics*》(New York: Hawthorn, 1970), Dirk J. Struik，《*A Concise History of Mathematics*》(New York: Wiley, 1989), Howard Eves，《*An Introduction to the History of Mathematics*》(Philadelphia: Saunders, 1990) 还有 David Burton，《*The History of Mathematics: An Introduction*》(Dubuque, Ia.: William C. Brown, 1991). 如果懂德文，应用参看 F. Klein 与其他人合编的《*Encyklopädie der mathematischen Wissenschaften*》(Leipzig: Teubner, 1898—1935)，一部浩大的著作，包含了数学科学各方面的专家们的长篇论述.

对原始文献，有许多从重要数学著作构成的选本，它们都被翻译成英语. 这些包括了 Ronakd Calinger 编纂的《*Classics of Mathematics*》(Englewood Cliffs, NJ: Prentice Hall, 1995), D. J. Struik 编纂的《*A Source Book in Mathematics, 1200—1800*》(Cambridge: Harvard University Press, 1969), Garrett Birkhoff 编纂的《*A Source Book in Classical Analysis*》(Cambridge: Harvard University Press, 1973), David Eugene Smith 编纂的《*A Source Book in Mathematics*》(New York: Dover, 1959)，还有 John Fauvel 和 Jeremy Gray 编纂的《*The History of Mathematics: A Reader*》(London: Macmillan,

1987).

当然,在数学史方面还继续有更多的研究出现,而且有许多发表此学科方面文章的刊物.可以在绝大多数大学图书馆中找到,最重要的这类刊物是《*Historia Mathematica*》和《*Archive for History of Exact Sciences*》.前者在每期中都发表一份数学史最新文章摘要的一览表.要紧跟现代文献最好还是去阅读《*Mathematical Reviews*》,由美国数学会每月出版,或者去参看《*Isis Current Bibliography of the History of Science and its cultural Influences*》,是 Isis 即每年出的科学史协会杂志的第5期.后面的期刊包含有按学科列出的前十二个月中有关科学史方面所发表文章的目录,当然也包括了数学史.这些资料来源今天可在许多研究性图书馆上联网供查.